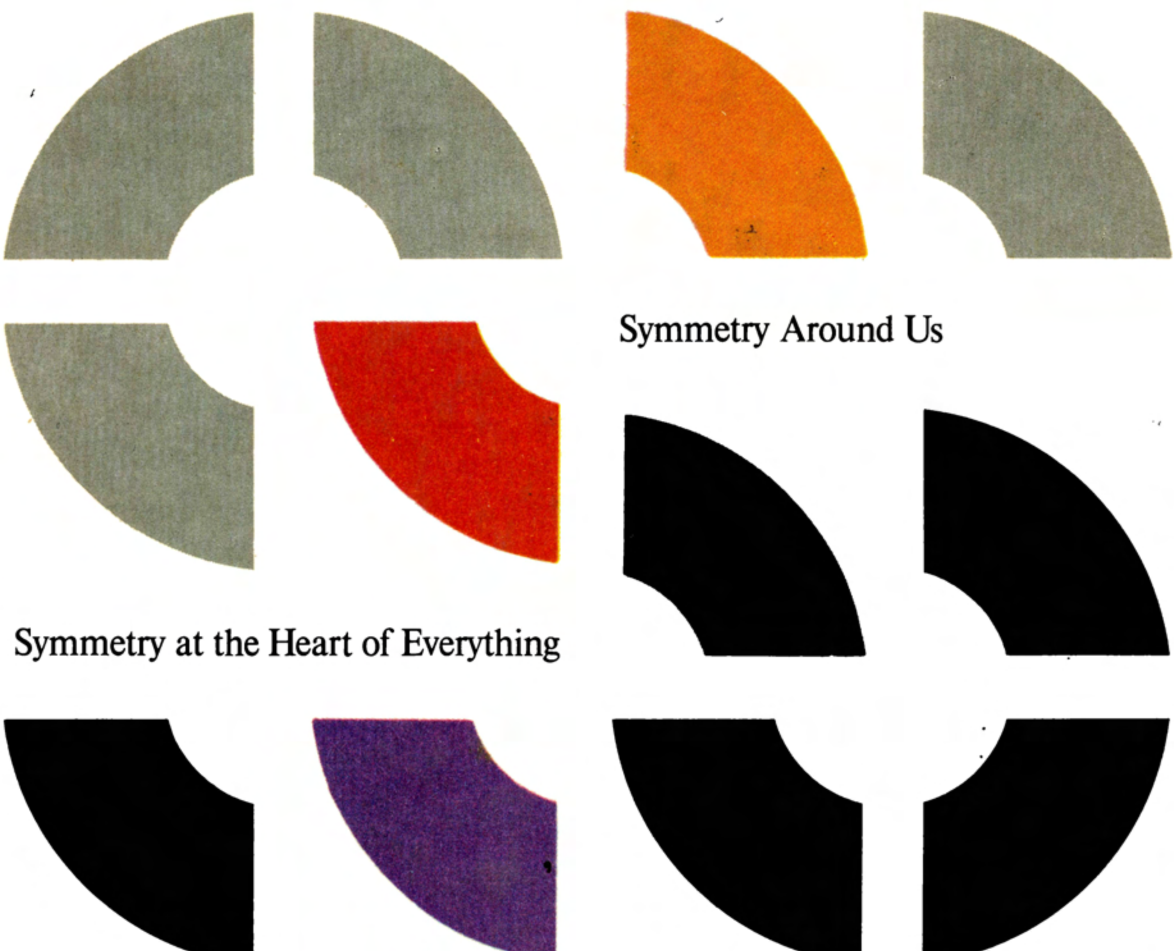
This Amazingly
Symmetrical World

L. Tarasov



Mir Publishers

About the Book

Here is a fascinating and nontechnical introduction to the ubiquitous effects of symmetry. Starting with geometrical elements of symmetry, the author reveals the beauty of symmetry in the laws of physics, biology, astronomy, and chemistry. He shows that symmetry and asymmetry form the foundation of relativity, quantum mechanics, solid-state physics, and atomic, subatomic, and elementary particle physics, and so on.

The author would like to attract the reader's attention to the very idea of symmetry and help him to discern a wide variety of the manifestations of symmetry in the surrounding world, and, above all, to demonstrate the most important role played by symmetry in the scientific comprehension of the world and in human creative effort.

This Amazingly
Symmetrical
World

Л. Тарасов

Этот удивительно симметричный мир

«Просвещение»
Москва

L. Tarasov

This Amazingly Symmetrical World

Translated from Russian by
Alexander Repyev

Mir Publishers
Moscow

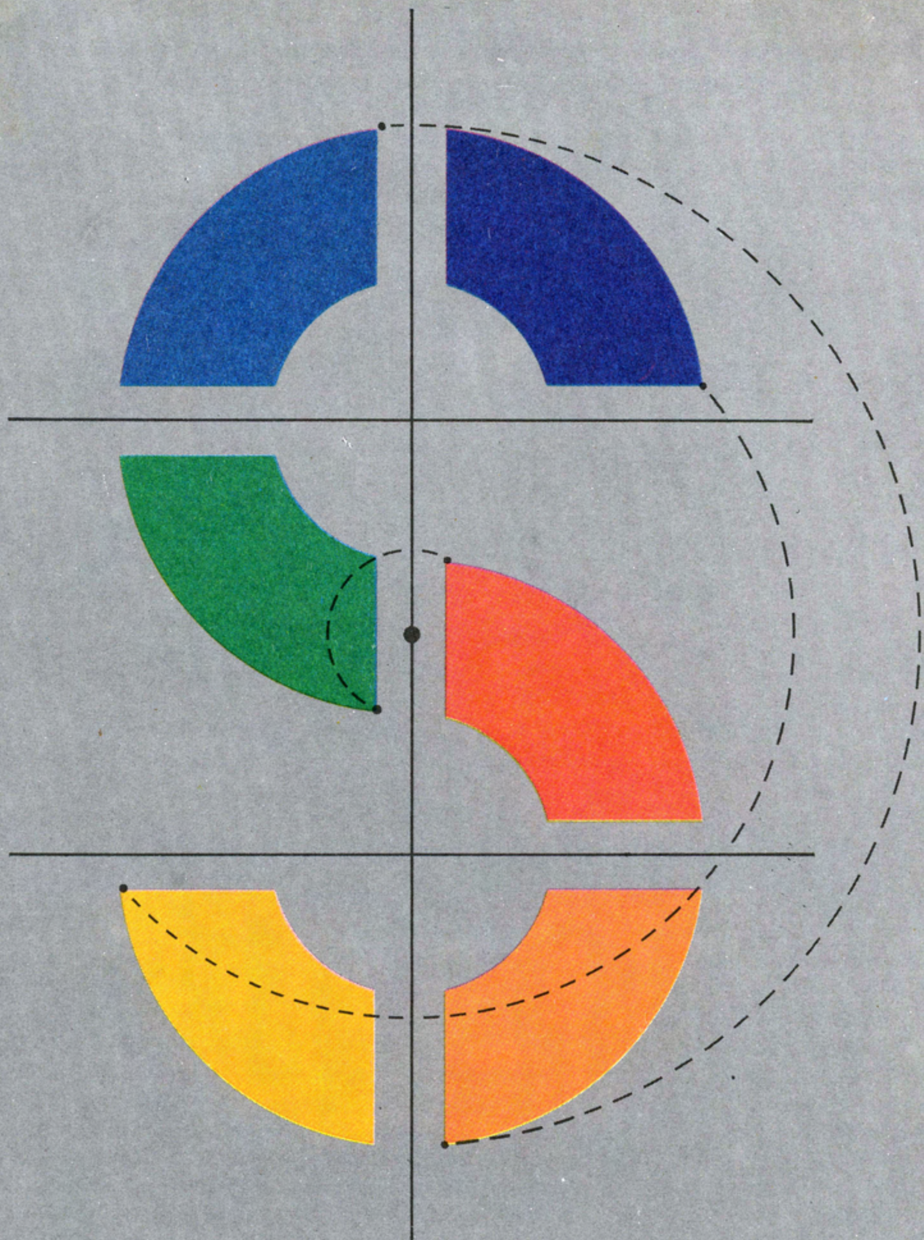
First published 1986

Revised from the 1982 Russian edition

На английском языке

Printed in the Union of Soviet Socialist Republics

Contents



Nature! Out of the simplest matter it creates most diverse things, without the slightest effort, with the greatest perfection, and on everything it casts sort of fine veil. Each of its creations has its own essence, each phenomenon has separate concept, but everything is a single whole.

Göthe

Preface

Symmetry is encountered everywhere – in nature, engineering, arts, and science. Note, for example, the symmetry of a butterfly and maple leaf, the symmetry of a car and plane, the symmetry of a verse and tune, the symmetry of patterns and borders, the symmetry of the atomic structure of molecules and crystals.

The notion of symmetry can be traced down through the entire history of human creative endeavours. It has its beginnings in the well-springs of human knowledge and it has widely been used by all the modern sciences. So principles of symmetry dominate in physics and mathematics, chemistry and biology, engineering and architecture, painting and sculpture, poetry and music. The laws of nature, which govern the infinite variety of phenomena, in turn obey the principles of symmetry.

Now then, what is symmetry? What profound idea underlies the concept? Why does symmetry literally permeate the entire world around us? Answers to these questions can be found in this book.

The book is arranged in two parts. The first part considers the symmetry of positions, forms, and structures. This is the sort of symmetry that can be seen directly. And so it can be thought of as geometric symmetry. The second part is concerned with the symmetry of physical phenomena and laws of nature. This symmetry lies at the very foundation of the naturalistic picture of the world, and so it can be called physical symmetry.

The book is written in simple, nontechnical language and is designed for the reader with but an elementary understanding of physics and mathematics. On the other hand, the book is not intended to be an easy, entertaining read. The reader will have to be patient at times to grasp some relatively difficult points, especially in the first chapters, which treat the notions of mirror, rotational, translational, and other kinds of symmetry. No special background is required here, just a measure of patience to be rewarded, as I hope, by the satisfaction derived from reading the subsequent chapters.

I would like to attract the reader's attention to the very idea of symmetry and help him to discern a wide variety of the manifestations of symmetry in the surrounding world, and above all, to demonstrate the most important role played by symmetry in the scientific comprehension of the world and in human creative effort.

The idea of writing the book was prompted by I. Gurevich, who was also very helpful in selecting the material and format of the book, for which I am grateful to him. Thanks are also due to A. Tarasova for her help with the manuscript.

L. Tarasov

A Conversation Between the Author and the Reader About What Symmetry Is

Standing at the blackboard and drawing some figures on it with chalk I was suddenly struck by the idea: why is symmetry so pleasing to the eye? What is symmetry? It is an innate feeling, I answered myself. But what is it based on?

Lev Tolstoy

AUTHOR. There is an old parable about the 'Buridan's ass'. The ass starved to death because he could not decide on which heap of food to start with (Fig. 1). The allegory of the ass is, of course, a joke. But take a look at the two balanced pans in the figure. Do the pans not remind us in some way of the parable?

READER. Really. In both cases the left and the right are so identical that neither can be given preference.

AUTHOR. In other words, in both cases we have a *symmetry*, which manifests itself in the total equity and balance of the *left* and *right*. And now tell me what you see in Fig. 2.

READER. The foreground shows a pyramid. Such pyramids were erected in Ancient Egypt. In the background are hills.

AUTHOR. And maybe in the foreground you see a hill as well?

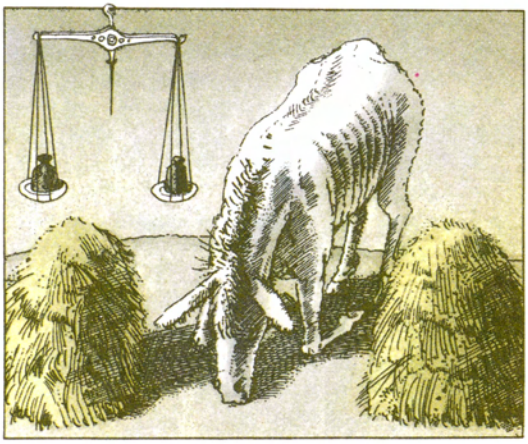


Fig. 1

READER. This is no hill. This looks like an artificial structure. A hill normally does not possess that regular, symmetric shape.
AUTHOR. Quite so. Perhaps you could provide quite a few examples of regular configuration (symmetry) of objects or structures created by man?

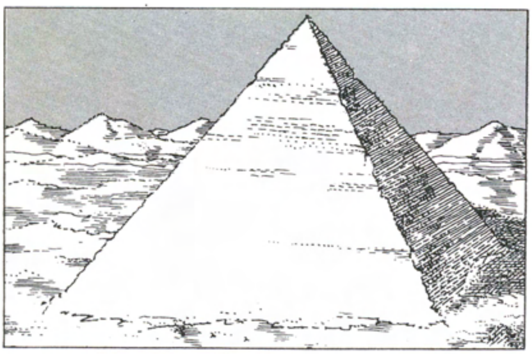


Fig. 2

READER. A legion of them. To begin with, architectural structures. For example, the building of the Bolshoi Theatre in Moscow. Essentially all vehicles, from a cart to a jet liner. Household utensils (furniture, plates, etc.). Some musical instruments: a guitar, a violin, a drum, ...
AUTHOR. Man-made things are often symmetrical in shape. Although this is not always the case, just remember a piano or a harp. But why do you think symmetry is so often present in human products?
READER. The symmetry of shape of a given object may be determined by purpose. Nobody wants a lop-sided ship or a plane with wings of different lengths. Besides, symmetrical objects are beautiful.
AUTHOR. That reminds me of the words of the German mathematician Hermann Weyl (1885-1955), who said that through symmetry man always tried to 'perceive and create order, beauty and perfection'.
READER. The idea of symmetry, it seems, is inherent in all spheres of human endeavour.
AUTHOR. Exactly. It would be erroneous, however, to think that symmetry is mostly present in human creations whereas nature prefers to appear in nonsymmetrical (or *asymmetrical*) forms. Symmetry occurs in nature no rarer than in man-made things. Since the earliest times nature taught man to understand symmetry and then use it. Who has not been moved by the symmetry of snow-flakes, crystals, leaves, and flowers? Symmetrical are animals and plants. Symmetrical is the human body.
READER. Really, symmetrical objects seem to surround us.
AUTHOR. Not just objects. Symmetry can also be seen in the regularity of the alternation of day and night, of seasons. It is present in the rhythm of a poem. Essentially we deal with symmetry everywhere where we have a *measure* of order. Symmetry, viewed in the widest sense, is the opposite of chaos and disorder.
READER. It turns out tnat symmetry is balance, order, beauty, perfection, and even purpose. But such a definition appears to be a bit too general and blurry, doesn't it? What then is meant by the term 'symmetry' more specifically? What features signal the presence, or absence for that matter, of symmetry in a given instance?
AUTHOR. The term 'symmetry' (σνμμετρυα) is the Greek for 'proportionality, similarity in arrangement of parts'.
READER. But this definition is again not concrete.
AUTHOR. Right. A mathematically rigorous concept of symmetry took shape only recently, in the 19th century. The simplest rendering (according to Weyl) of the modern definition of symmetry is as follows: *an object is called symmetrical if it can be changed somehow to obtain the same object*.
READER. If I understand you correctly, the modern conception of symmetry assumes the *unchangeability* of an object subject to some *transformations*.
AUTHOR. Exactly.
READER. Will you please clarify this by an example.

AUTHOR. Take, for example, the symmetry of the letters U, H and N.
READER. But N does not appear to be symmetrical at all.
AUTHOR. Let's start with U. If one half of the letter were reflected in a plane mirror, as is shown in Fig. 3*a*, the reflection would coincide exactly with the other half. This is an example of so-called *mirror symmetry*, or rather a symmetry in relation to mirror reflection. We have already encountered this form of symmetry: it then appeared as an equilibrium of the left and right sides (the left and the right in Fig. 1 can be viewed as mirror reflections of each other).

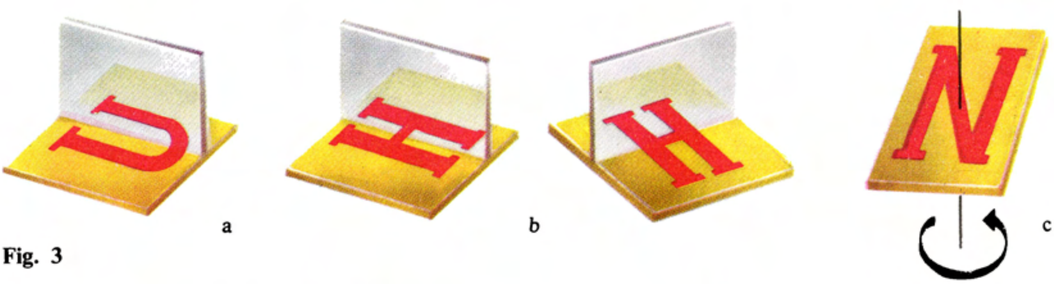


Fig. 3

Letter H is symmetrical to an even higher degree than U. It can be reflected in two plane mirrors (Fig. 3*b*). As for the letter N, it has no mirror symmetry, but it has the so-called *rotational symmetry* instead. If N were turned through 180° about the axis, that is normal to the letter's plane and passes through its centre (Fig. 3*c*), the letter would coincide with itself. Put another way, the letter N is symmetrical in relation to a turn through 180°. Note that rotational symmetry is also inherent in H, but not in U.
READER. An Egyptian pyramid too possesses a rotational symmetry. It would coincide with itself if mentally turned through 90° about the vertical axis passing through the pyramid's summit.
AUTHOR. Right. In addition, a pyramid possesses a mirror symmetry. It would coincide with itself if mentally reflected in any of the four imaginary planes shown in Fig. 4.
READER. But how is one to understand the above-mentioned symmetry in the alternation of seasons?
AUTHOR. As the unchangeability of a certain set of phenomena (including weather, blooming of plants, the coming of snow, etc.) in relation to shift in time over 12 months.

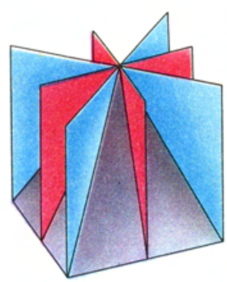


Fig. 4

READER. The beauty of symmetrical objects and phenomena seems to come from their harmony and regularity.
AUTHOR. The question of the beauty related to symmetry is not that obvious.
READER. Why not? I can imagine that looking at harmonious, balanced, recurrent parts of a symmetrical object must give one the feeling of peace, stability, and order. As a result, the object will be perceived as beautiful. And even more so, if in its symmetry we also see some purpose. On the contrary, any accidental violation of symmetry (the collapsed corner of a building, the broken piece of a neon letter, an early snow) must be perceived negatively – as a threat to our trust

in the stability and orderliness of the surrounding world.

AUTHOR. That is all very well but it is known that symmetry may also produce negative emotions. Just look at some modern residential areas consisting of identical symmetrical houses (often fairly convenient and rational), do they not bore you to death? On the other hand, some deviations from symmetry used widely in painting and sculpture create the mood of freedom and nonchalance, and impart an inimitable individuality to a work of art. Hardly anyone would be immune to the spell of a spring meadow in full bloom with an absolutely asymmetrical array of colours. Do you really think that a neatly mown lawn or a trimmed tree look prettier than a clearing in a forest or an oak growing in a field?

READER. Beauty is thus not always related to symmetry.

AUTHOR. The truth is that when considering symmetry, you have to take into account not only symmetry as such, but also deviations from it. *Symmetry* and *asymmetry* must be approached simultaneously.

READER. Perhaps as is the case in nature?

AUTHOR. Of course. But one important point here: one should not merely look at given violations of symmetry in a specific flower or organism of an animal. The issue of symmetry-asymmetry is far deeper. Symmetry may be said to express something *general*, characteristic of different objects and phenomena, it is associated with the *structure* and lies at the very foundation of things. Asymmetry, on the other hand, is related to realizations of the structure of a given *specific* object or phenomenon.

READER. Symmetry is general, asymmetry is specific?

AUTHOR. If you like, you may think of it this way. In a specific object we find elements of symmetry, which link it to other similar objects. But the individual 'face' of a given object invariably shows up through the presence of some element of asymmetry. Spruces have much in common: a vertical stem, characteristic branches arranged in a certain rotational symmetry around the stem, a definite alternation of branches along the stem, and lastly the structure of needles. And still you take your time selecting the right spruce at a Christmas tree bazar. Among the many trees available you look for individual traits that you like most.

READER. It turns out then that the mathematical idea of symmetry in each case is embodied in real, not-very-symmetrical objects and phenomena.

AUTHOR. Let's try and visualize a world, arranged in a totally symmetrical manner. Such a world would have to be able to be reproduced in any rotation about any axis, in the reflection of any mirror, in any translation, and so on. This would be something absolutely homogeneous, uniformly spread out over the entire space.

Stated another way, in that sort of world you would observe nothing, neither objects nor phenomena. And so such a world is impossible.

The world exists owing to the marriage of symmetry and asymmetry, which can in a way be treated as the unity of the general and the specific.

READER. Frankly, I never thought of symmetry in such a broad context.

AUTHOR. Let's draw some conclusions. Symmetry occurs widely, both in nature and in human life. Therefore, even a layman generally has no difficulty in discerning symmetry in its relatively simple manifestations.

Our world, all the things and phenomena in it, must be approached as a manifestation of symmetry and asymmetry. In that case, symmetry is not just abundant, it is ubiquitous, in the broadest sense.

Symmetry is utterly diverse. Objects may remain unchanged under turns, reflections, translations, interchanges, and so forth.

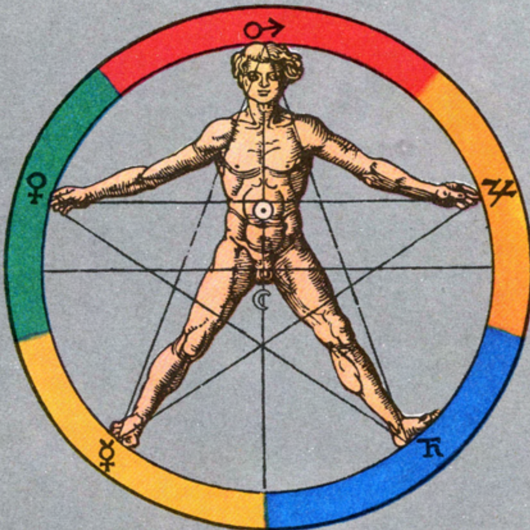
Symmetry has many causes. It may be related to orderliness and equilibrium, proportionality and harmony of parts (and sometimes monotony), purposefulness and usefulness.

Part ONE

Symmetry Around Us

Tyger! Tyger! burning bright
In the forests of the night,
What immortal hand or eye
Dare frame thy fearful symmetry?

W. Blake



1 Mirror Symmetry

What can be more like my hand or my ear than their reflections in a mirror? And still the hand I see in the mirror cannot substitute for my real hand.

I. Kant

An Object and Its Mirror Twin

Figure 5 shows a simple example of an object and its '*looking-glass twin*' – triangle ABC and triangle $A_1B_1C_1$ (here MN is the intersection of the mirror plane with the plane of the drawing). For each point on the triangle there is a corresponding point on the mirror twin. These points lie on the same perpendicular to straight line MN, on either side and at the same distance from it. For simplicity, the object in the figure is taken to be two-dimensional. In the general case, the object (and hence its mirror twin) are three-dimensional.

We all know that a mirror twin can be seen readily. Suffice it to place an illuminated object in front of a plane mirror and look into it. It is generally believed that the twin in the mirror is an exact replica of the object itself. But this is actually not so. A mirror does not just duplicate the object but transposes the front and back (in relation to the mirror) parts of the object. In comparison with the object itself, its mirror twin seems to be turned inside out in the direction perpendicular to the plane of the mirror. This is clearly seen in Fig. 6*a* and essentially imperceptible in Fig. 6*b*. When examining the cones in the figure the reader might disagree with our statement that the mirror reflection is not an exact copy of the object. After all, the object and its reflection in Fig. 6*a* only differ in their orientation: they are so arranged that they face each other. This is all the more so for Fig. 6*b*. In this connection let us turn to a more interesting example.

Suppose that a cone is rotated about its axis (Fig. 7), the direction of rotation being indicated by the arrow. If the rotation axis is normal to the mirror's plane, the cone retains its direction of rotation when reflected in the mirror (Fig. 7*a*). If the rotation axis is parallel to the mirror, the direction of rotation is reversed in the reflection (Fig. 7*b*). No shifts or turns can now make the object coincide with its reflection. In other words, the rotating cone and its reflection are essentially *different* objects. To obtain a reflection without actually reflecting the cone in the mirror, we will have to reverse the direction of the rotation of the cone.

We can do this without any rotation, though. If we make a *thread* on the cone (Fig. 8), then the object's thread and the reflection's thread will have different directions: for the object's thread to be driven into wood we will have to rotate its head clockwise, and the

reflection's one counterclockwise. The first thread is called *right*, and the second *left*. We generally use right screws.

We have thus seen that, whatever the amount of similarity, an object and its reflection may be different, *incompatible*. Sometimes the difference is not very conspicuous; for instance, you generally attach no importance to the fact that a birth-mark on your right cheek in your 'looking-glass' counterpart appears on the left cheek. In other cases the difference becomes so striking that we can only be surprised

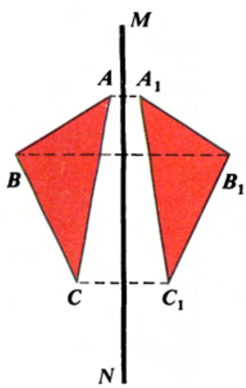


Fig. 5

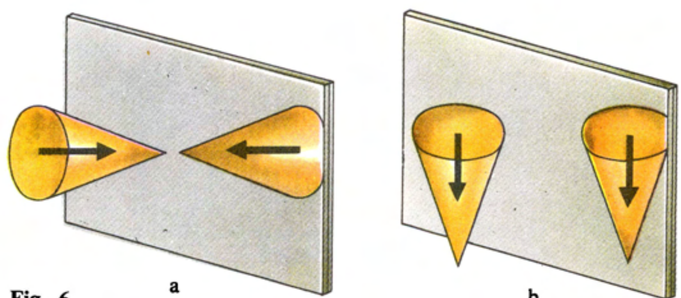


Fig. 6

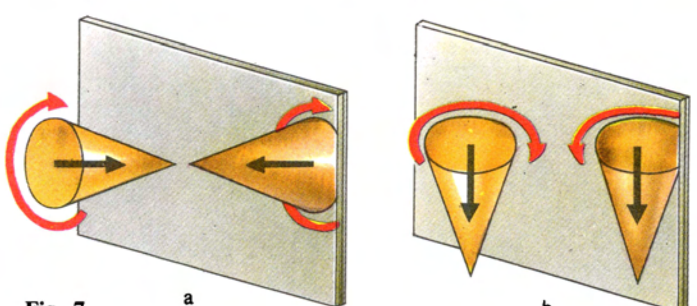


Fig. 7

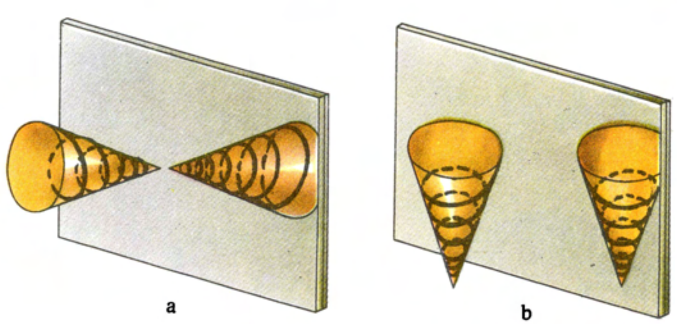


Fig. 8

that it has been overlooked earlier. It is sufficient to compare a text with its reflection (Fig. 9). Hold a book up to a mirror and try and read the reflection of the text in it. Or, worse, try and write just one line looking not at the paper but at its mirror reflection.

Mirror Symmetry

Suppose now that one half of an object is a mirror reflection of the other about a plane. Such an object is said to be *mirror-symmetrical*,

Fig. 9

and the plane is said to be the *plane of symmetry*. For a two-dimensional (flat) object instead of a plane of symmetry we have an *axis of symmetry*–the line of intersection of the plane of symmetry with the plane of the object. For a unidimensional (linear) object we have the *centre of symmetry*–the point of the intersection of the object's straight line with the plane of symmetry.

Figure 10 gives examples of mirror-symmetrical objects: (a) unidimensional object (O is the centre of symmetry), (b) two-dimensional object (MN is the axis of symmetry), (c) three-dimensional object (S is the plane of symmetry).

A unidimensional object has no more than one centre of symmetry. A two-dimensional object may have several axes of symmetry, and

a three-dimensional object, several planes of symmetry. So a regular hexagon has six axes of symmetry (the red lines in Fig. 11). In Fig. 4 were shown four planes of symmetry of a regular pyramid. The circle has an infinite number of axes of symmetry, just as the sphere, the circular cylinder, the circular cone, and the spheroid.

Let us print on a sheet of paper in capital letters the words COOK and PAN. We will now get a mirror and place it vertically so that the line of the intersection of the mirror's plane with the sheet's plane

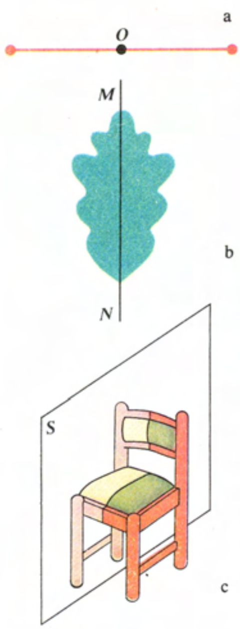


Fig. 10

Fig. 12

divides these words in two along the horizontal line. Perhaps some will be surprised at finding that the mirror has not changed the word COOK, whereas it has changed the word PAN beyond recognition (Fig. 12). This 'trick' can be explained very simply. To be sure, the mirror reflects the lower part of both words in a similar fashion. Unlike PAN, however, the word COOK possesses a horizontal axis of symmetry, that is why it is not distorted by the reflection.

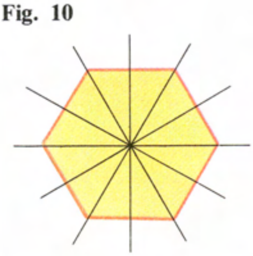


Fig. 11

Enantiomorphs

Suppose that an object is characterized by a *single* plane (axis) of symmetry. Let us cut it along the plane (axis) of symmetry into two halves. These two halves are, clearly, mirror reflections of each other. It is essential that each half in itself is mirror-asymmetrical. The halves are said to be *enantiomorphs.*

Thus, *enantiomorphs are pairs of mirror-asymmetrical objects (figures) that are mirror reflections of each other.* In other words, enantiomorphs are an object and its mirror reflection provided that the object itself is mirror-asymmetrical. Enantiomorphs may be individual objects and halves of an object cut appropriately. To distinguish between enantiomorphs of a given pair they are referred to as *left* and *right.* It is immaterial which is called left (right), it is only a matter of convention, tradition, or habit.

Examples of three-dimensional enantiomorphs are given in Fig. 13: (a) left and right screws, (b) left and right dice, (c) left and right knots, (d) left and right gloves, (e) left and right reference frames, (f) left and right halves of a chair cut along the symmetry plane. Normally used in practice are right screws, left dice, and right reference frames. Left and right gloves and knots occur with equal frequency.

Given in Fig. 14 are examples of two-dimensional enantiomorphs: (a) left and right spirals, (b) left and right traffic signs, (c) left and right

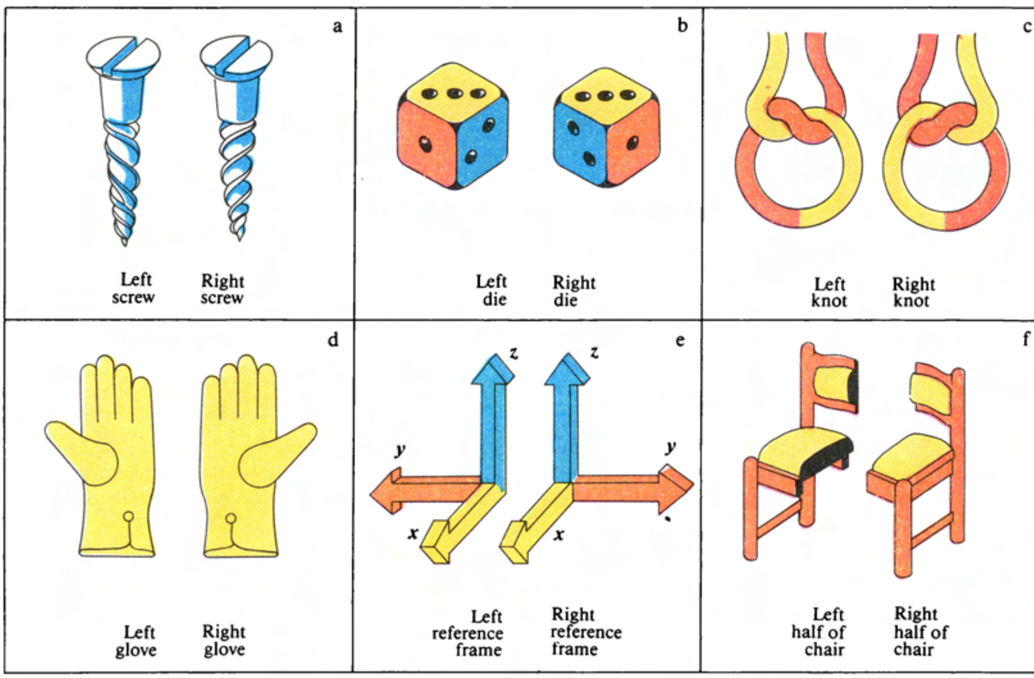


Fig. 13

reference frames, (d) left and right halves of an oak leave cut along the axis of symmetry.

Two-dimensional enantiomorphs cannot be superposed on each other by any translations or turns in the space of these enantiomorphs, that is, in a plane. To make them coincide a turn in the three-dimensional space is required, that is, a turn as shown in Fig. 15. As for *three-dimensional* enantiomorphs, for them to coincide exactly, turning in a fantastic non-existent four-dimensional space would be required. It is clear that to effect the turn or even to visualize it is impossible. Therefore, for three-dimensional enantiomorphs the following statement is valid: *no translocations or turns can convert a left enantiomorph into its right one, and vice versa.* So a left enantiomorph will always be left and a right one will always be right. Turn your left

shoe as you might, it will never fit your right foot. Cast a left die as often as you might, it will never turn into a right die.

Curiously, to prove the existence of an 'other-worldly' four-dimensional world they contrived tricky demonstrations involving conversions of left enantiomorphs into right ones, and vice versa. Such demonstrations were performed at so-called spiritualistic sessions, which were rather fashionable at the turn of the century in some

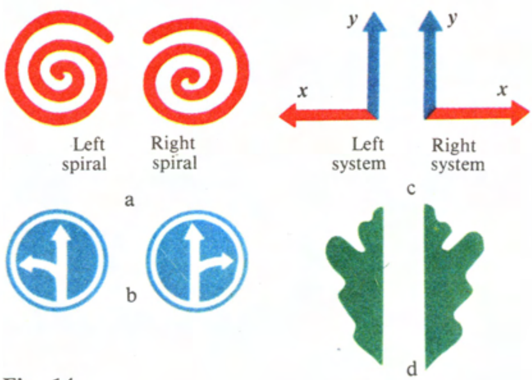


Fig. 14

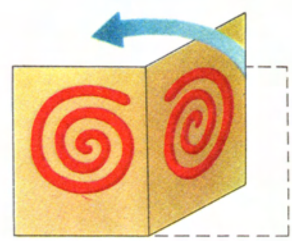


Fig. 15

decadent aristocratic circles. It goes without saying that those demonstrations were just adroit tricks based on sleight of hand. For example, the spiritualist would ask a participant in the session to give him his left glove, and after some distracting manipulations, the spiritualist would present to the spectators sitting in a half-dark room an identical right glove. This was claimed to be a proof of the short-term residence of the glove in the other world, where it had allegedly turned into the right glove.

2 Other Kinds of Symmetry

If we venture beyond our habitual notion of symmetry as a property that is in all ways related to our external appearances, we can find quite a few figures that are symmetrical in one or another sense.

A. S. Kompaneets

Rotational Symmetry

Let there be an object that coincides with itself exactly when turned about some axis through an angle of $360°/n$ (or its multiple), where $n = 2, 3, 4, \ldots$. In that case, we speak about *rotational symmetry*, and the above-mentioned axis is referred to as the *n-fold rotation symmetry axis*, or *n-fold symmetry*, or *axis of n-fold symmetry*. In the earlier examples of letters H and N we dealt with a two-fold axis, and in the example with an Egyptian pyramid, with a 4-fold axis. Figure 16 gives examples of simple objects with rotation axes of different orders – from two-fold to five-fold.

A three-dimensional object may have several rotation axes. For example, the first object in Fig. 16 has three two-fold axes, the second object, in addition to a three-fold axis, has three two-fold axes, and the third object, in addition to a four-fold axis, has four two-fold axes (additional axes are shown in the figure by dash lines).

Consider the *cube*. We can see immediately that it has three four-fold rotation axes (Fig. 17*a*). Further examination reveals six two-fold axes passing through the centres of opposite parallel edges (Fig. 17*b*), and four three-fold axes, which coincide with the inner diagonals of the cube (Fig. 17*c*). The cube thus has 13 rotation axes in all, among them two-, three-, and four-fold ones.

Also interesting is the rotational symmetry of the *circular cylinder*. It has an infinite number of two-fold axes and one infinite-fold axis (Fig. 18).

To describe the symmetry of a concrete object requires specifying all the axes and their orders, and also all the planes of symmetry. Let us take, for example, a geometrical body composed of two identical regular pyramids (Fig. 19). It has one four-fold rotation axis (*AB*), four two-fold axes (*CE, DF, MP*, and *NQ*), five planes of symmetry (*CDEF, AFBD, ACBE, AMBP*, and *ANBQ*).

Mirror-Rotational Symmetry

Let us cut a square out of thick paper and inscribe into it obliquely another square (Fig. 20). Now we will bend the corners along the periphery of the inner square as shown in Fig. 21. The resultant object will have a two-fold symmetry axis (*AB*) and no planes of symmetry. We will view the object first from above and then from below (from the other side of the sheet). We will find that the top and bottom views

look the same. This suggests that the two-fold rotational symmetry does not exhaust all the symmetry of a given object.

The additional symmetry of this object is the so-called *mirror-rotational symmetry*, in which the object coincides with itself when turned through 90° about axis AB and then reflected from plane $CDEF$. Axis AB is called the four-fold *mirror-rotational axis*. We thus have a symmetry relative to two consecutive operations – a turn by 90° and a reflection in a plane normal to the rotation axis.

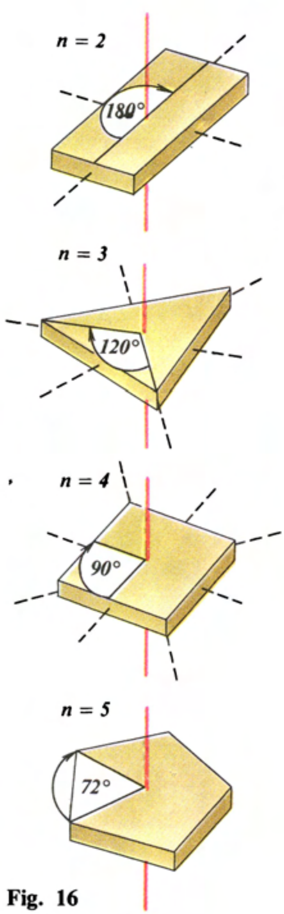


Fig. 16

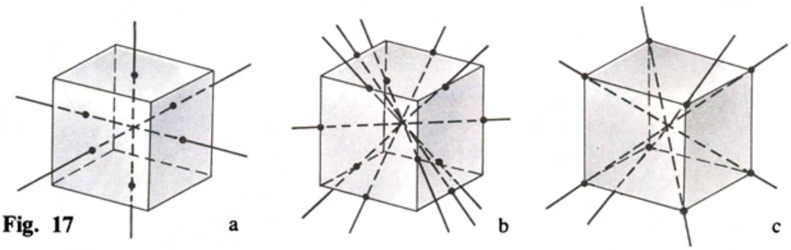


Fig. 17

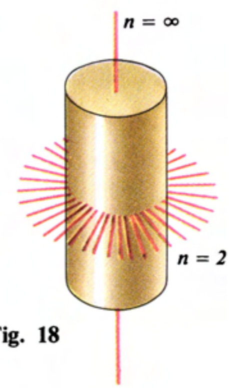


Fig. 18

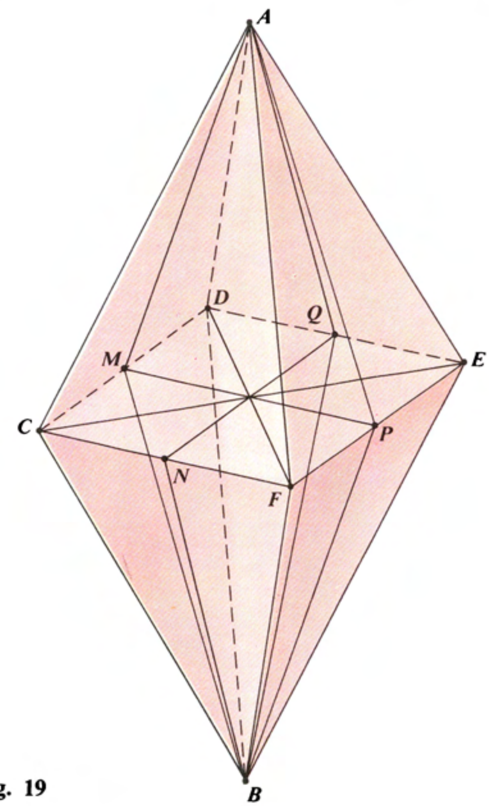


Fig. 19

Translational Symmetry

Consider the plane figure in Fig. 22*a*. A translation along line *AB* by a distance *a* (or its multiple) makes the figure seem unchanged. In that case we speak about *translational symmetry*. *AB* is called the *axis of translation*, and *a* is called the *fundamental translation*, or *period*. Strictly speaking, a body having a translational symmetry must be infinite in the direction of the translation. But the concept of translational symmetry in a number of cases is also applied to bodies of finite sizes when partial coincidence is observed. It is seen in Fig. 22*b* that when a finite figure is displaced by distance *a* along line *AB*, part *1* coincides with part *2*.

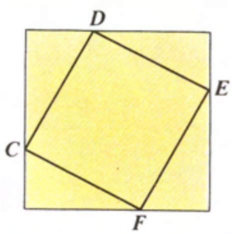


Fig. 20

Associated with translational symmetry is the important concept of a two-dimensional *periodic structure* – the *plane lattice*. A plane lattice can be formed by crossing two sets of parallel equispaced straight lines (Fig. 23), their intersections being called *sites*. To specify a lattice, it is sufficient to specify its *unit cell* and then to displace it along *AB* by distances multiple of *a*, or along *AC* by distances multiple of *b*. Note that in a given lattice a unit cell may be chosen in a *wide variety of ways*. So we may select the red cell in Fig. 23. Or we may also use any of the shaded ones.

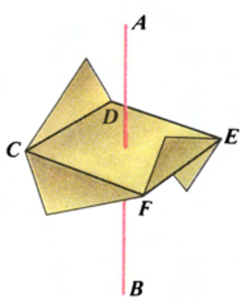


Fig. 21

The translational symmetry of a plane lattice is totally defined by a combination of two vectors ($\vec{a}$ and $\vec{b}$ in Fig. 23). Five types of plane lattices (five types of translational symmetry in a plane) are distinguished. These are given in Fig. 24: (a) $a = b$, $\gamma = 90°$ (*square* lattice), (b) $a \neq b$, $\gamma = 90°$ (*rectangular* lattice), (c) $a = b$, $\gamma = 60°$ (*hexagonal* lattice), (d) $a = b$, $\gamma \neq 90°$, $\gamma \neq 60°$ (*rhombic* lattice), and (e) $a \neq b$, $\gamma \neq 90°$ (*oblique* lattice).

Translational symmetry in three-dimensional space is associated with the notion of a three-dimensional periodic structure – the *space lattice*. This lattice can be thought of as a result of the crossing of three sets of parallel planes. The translational symmetry of a three-dimensional lattice is defined by the three vectors that specify the unit cell of the lattice. Figure 25 shows a unit cell described by $\vec{a}$, $\vec{b}$, and $\vec{c}$. In the simplest case, all the edges of a cell are equal in length and the angles between them are 90°. In that case, we have a *cubic lattice*. All in all, there are 14 types of space lattices, differing in their type of space symmetry. In other words, there are 14 *Bravais lattices* (named after a French crystallographer of the 19th century).

Bad Neighbours

Translational and rotational symmetries may live side by side with each other. So, the square lattice (Fig. 24*a*) has a four-fold rotational symmetry, and the hexagonal lattice (Fig. 24*c*) has a six-fold rotational symmetry. It is easily seen that a lattice has an infinite number of rotation axes. For example, in the case of the square lattice, the rotation axes (four-fold) pass through the centre of each square cell and through each lattice site.

But translational and rotational symmetries are bad neighbours. *In*

the presence of a translational symmetry only two-, three-, four-, and six-fold rotation axes are possible. Let's prove this.

Let points A and B in Fig. 26 be sites of some plane lattice ($|AB| = a$). Suppose that through these sites n-fold rotation axes pass perpendicular to the plane of the lattice. We will now turn the lattice about axis A through an angle $\alpha = 360°/n$, and will designate by C the new position of site B. If the lattice were turned through the angle α around axis B in the opposite direction, site A would come to D.

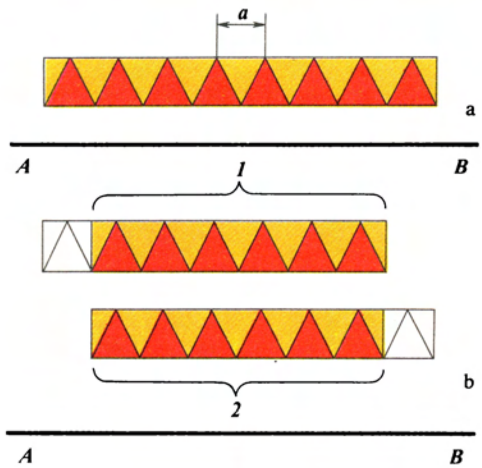


Fig. 22

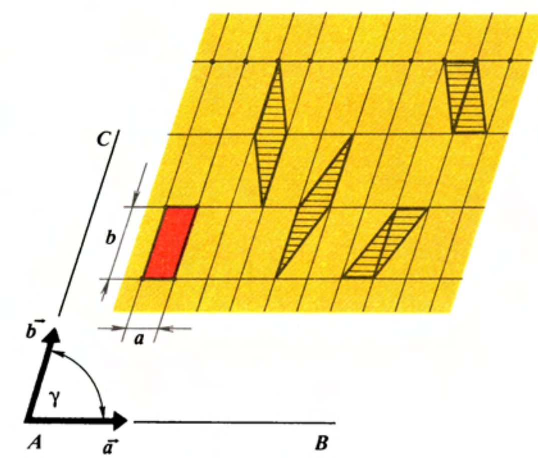


Fig. 23

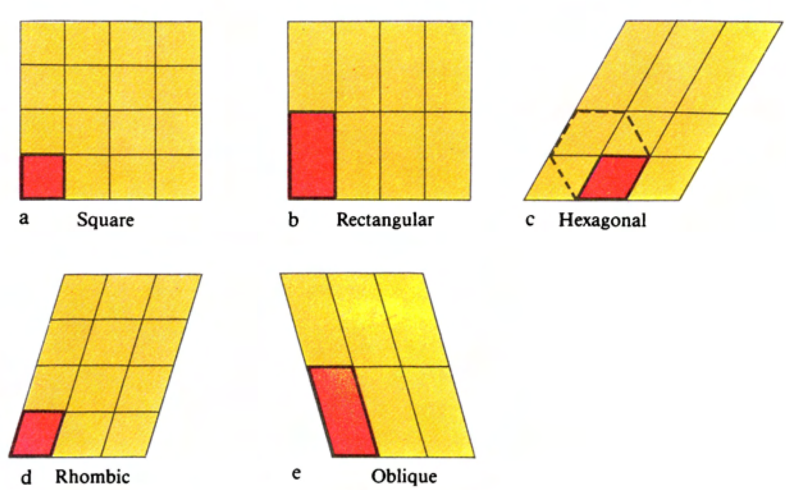


Fig. 24

The presence of translational symmetry requires that points C and D coincide with lattice sites. Hence,

$$|CD| = m|AB| = ma,$$

where m is an integer. The trapezoid $ABDC$ (see the figure) being equilateral, we have $|CD| = a \pm 2a\cos\alpha$. Thus,

$$a(1 \pm 2\cos\alpha) = ma,$$

that is, $\cos\alpha = \pm(m-1)/2$. Since $|\cos\alpha| \leqslant 1$,

$$-2 \leqslant (m-1) \leqslant 2.$$

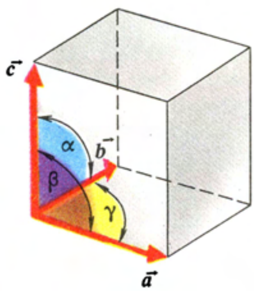


Fig. 25

It follows that only the following five cases are possible:

(1) $m = -1$, $\cos\ \alpha = -1$, $\alpha = 180°$, $n = 2$ (two-fold rotational symmetry);

(2) $m = 0$, $\cos\alpha = -1/2$, $\alpha = 120°$, $n = 3$ (three-fold rotational symmetry);

(3) $m = 1$, $\cos\ \alpha = 0$, $\alpha = 90°$, $n = 4$ (four-fold rotational symmetry);

(4) $m = 2$, $\cos\ \alpha = 1/2$, $\alpha = 60°$, $n = 6$ (six-fold rotational symmetry);

(5) $m = 3$, $\cos\alpha = 1$, $\alpha = 0$.

Accordingly, with translational symmetry, five-fold rotation axes are impossible in principle, as well as axes of an order higher than six.

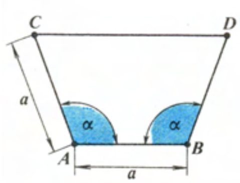


Fig. 26

Glide Plane (Axis) of Symmetry

It was shown earlier that successive turns and reflections may give rise to a new type of symmetry – the mirror-reflection symmetry. Combining turns or reflections with translations can also produce new types of symmetry. By way of example, consider a symmetry involving the so-called *glide plane of symmetry* (or rather *glide axis of symmetry*, since the figure in question is a plane). Figure 27 depicts a design exhibiting a translational symmetry along axis AB with period $2a$. It is easily seen that another symmetry can be revealed here – the symmetry relative to the displacement along AB with period a followed by the reflection about axis AB. Axis AB is termed a glide axis of symmetry with period a.

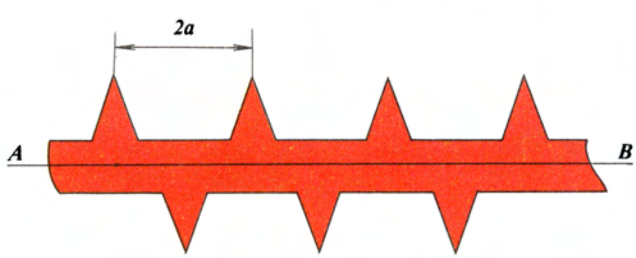


Fig. 27

3 Borders and Patterns

A mathematician, just like an artist or a poet, produces designs.

G. Hardy

Borders

A periodically recurring pattern on a tape is called a *border*. Borders come in various types. These may be frescoes, decorating walls, galleries, stairs, and also iron casting used as fencing for parks, bridges, embankments. These also may be gypsum plaster-reliefs or earthenware.

Figure 28 presents 14 borders in seven pairs. Each pair consists of borders similar in symmetry type. *Altogether, there are seven types of symmetry of borders.* Any border has a *translational* symmetry along its axis (translation axis). In the simplest case, the border has only translational symmetry (Fig. 28*a*). Figure 29*a* is a schematic representation of that sort of border, the triangle stands for a recurring asymmetric element.

The borders shown in Fig. 28*b*, apart from a translational symmetry, also have a mirror symmetry: they are mirror-symmetrical relative to the straight line dividing the border tape in half longitudinally. This sort of border is shown schematically in Fig. 29*b*, where the translation axis doubles as an axis of symmetry.

In the borders depicted in Figs. 28*c* and 29*c* the translation axis is a glide axis.

The borders in Fig. 28*d* have transverse axes of symmetry. They are given in Fig. 29*d* as segments of straight lines perpendicular to the translation axis.

The borders of Fig. 28*e* have two-fold rotation axes perpendicular to the border plane. The intersections of those axes with the border plane are marked in Fig. 29*e* by lentil-shaped figures.

Combining glide axes with two-fold rotation axes normal to the border plane produces the borders given in Fig. 28*f*, which possess transverse axes of symmetry. The scheme of that kind of border is presented in Fig. 29*f*.

Lastly, Figs. 28*g* and 29*g* provide borders based on combinations of mirror reflections. Apart from a longitudinal axis, such borders also have transverse axes of symmetry. As a consequence, two-fold rotation axes emerge.

Decorative Patterns

You have admired decorative patterns – those amazing designs that occur so widely in applied arts. In them you can find bizarre marriages

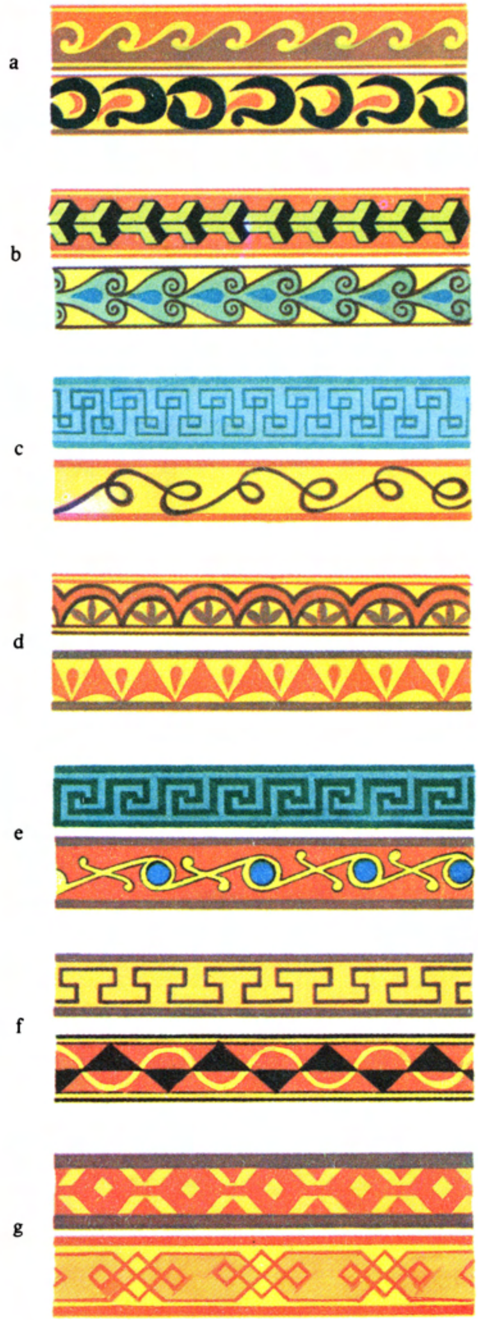


Fig. 28

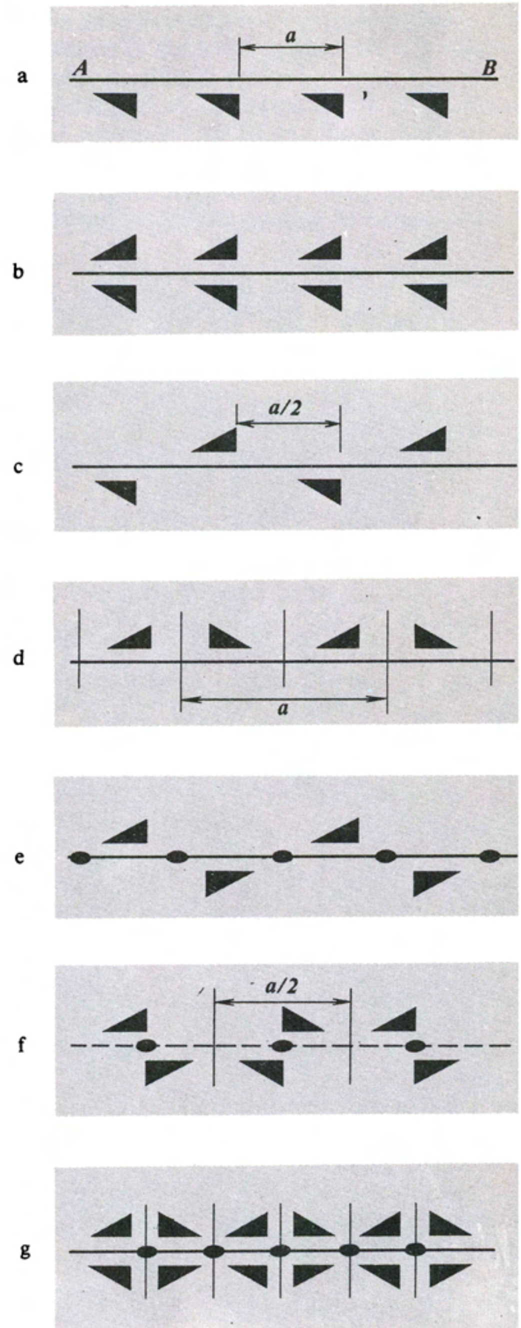


Fig. 29

of translational, mirror, and rotational symmetries. Examples abound. Just look at the design of the wall-paper in your room. Some patterns are shown in Figs. 30-32. Two of them have been produced by the eminent Dutch artist Escher: 'Flying Birds' (Fig. 30) and 'Lizards' (Fig. 32).

Any pattern is based on one of the five plane lattices discussed in Chapter 2. *The type of the plane lattice determines the character of the translational symmetry of a given pattern.* So the 'Flying Birds' pattern

Fig. 30

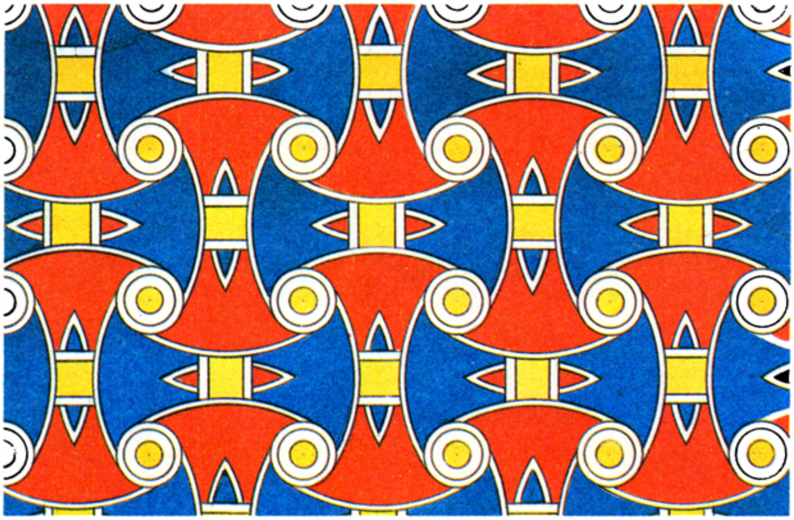


Fig. 31

is based on an *oblique* lattice, the characteristic Egyptian ornament in Fig. 31 on a *square* lattice, and the 'Lizards' pattern on a *hexagonal* lattice.

In the simplest case, a pattern is characterized by a *translational symmetry* alone. Such is, for example, the 'Flying Birds' pattern. To construct that pattern one must select an appropriate oblique lattice and 'fill' a unit cell of the lattice with some design, and then reproduce it repeatedly by displacing the cell without changing its orientation. In

Fig. 32

Fig. 33 a unit cell of the pattern is hatched. Note that the surface area of the cell is equal to the total area occupied by birds of different colours.

The symmetry of the Egyptian design is analyzed in Fig. 34. The translational symmetry of the pattern is given by a square lattice with a unit cell singled out in Fig. 34*a*. That cell has two-fold rotation axes, conventional and glide axes of symmetry. In Fig. 34*b* conventional axes of symmetry are indicated by solid lines, and glide axes by dash

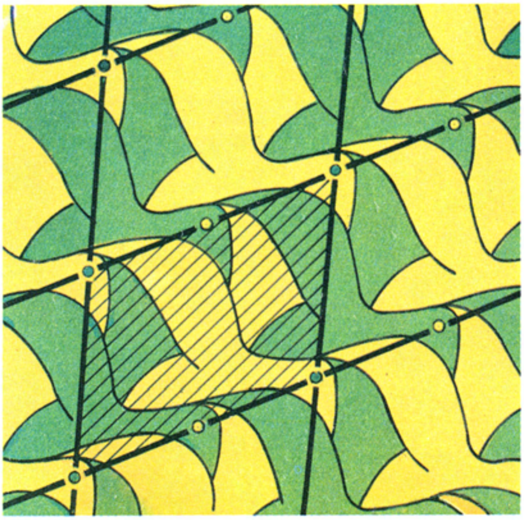


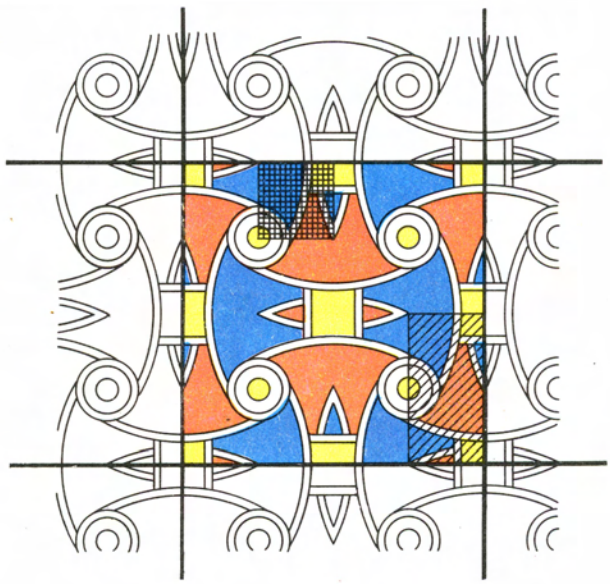
Fig. 33

lines. Intersections of the two-fold rotation axes with the plane of the pattern are marked by lentil-shaped figures. Unlike the 'Flying Birds', this design has a higher symmetry, as follows from the presence of *rotation* axes, and also of conventional and glide axes of *mirror symmetry*.

The symmetry of the Egyptian pattern would be yet higher if we simplified the colouring – instead of the red and blue colours used one colour, for example red. In that case, we would also have four-fold rotation axes and more glide axes of symmetry. The symmetry of this design is indicated in Fig. 34*c*, where the black squares are intersections of four-fold rotation axes with the plane of the pattern.

Figure 34*b* and *c* contains all the information about the elements of symmetry of the corresponding patterns. If in the figures we removed the design and only retained the conventional and glide axes and intersections of rotation axes with the plane of the pattern, we would obtain *schematic* representations of two different types of the symmetry of patterns. *In all, there are 17 types of symmetry of plane*

designs. They are given in Fig. 35. Here solid straight lines are conventional axes of symmetry and dash lines are glide axes. The lentils are the intersections of two-fold axes with the plane of the pattern; the triangles, three-fold axes; squares, four-fold axes; and hexagons, six-fold axes. The pattern of Fig. 34*b* is represented in Fig. 35 by #*9*, and the pattern of Fig. 34*c*, by #*12*; the 'Flying Birds', by #1.



a

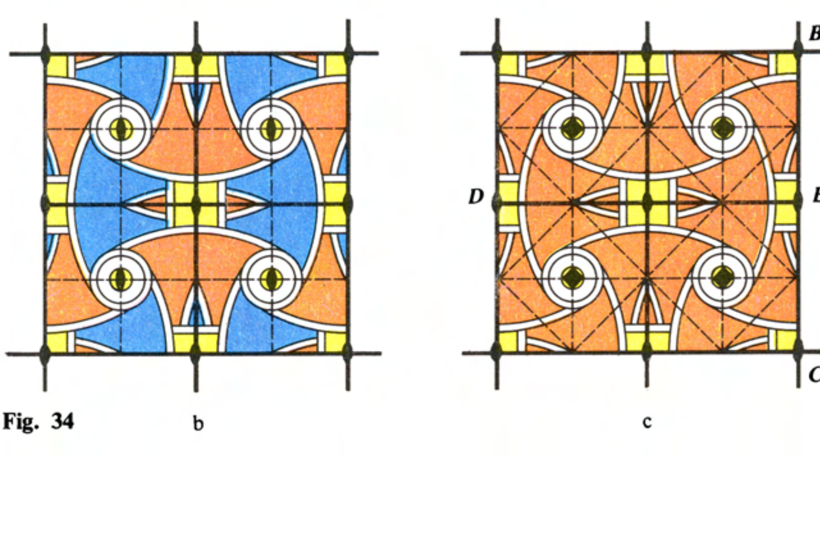


Fig. 34 b c

Pattern Construction

Any pattern can in principle be constructed along the lines of 'Flying Birds' by the parallel displacements of the unit cell, which is filled by some design. This procedure is the only one possible where a pattern has neither rotational nor mirror symmetry. Otherwise, a pattern can be produced using other procedures; the initial design then (or *main motif*) is not the entire unit cell of a design, but just part of it.

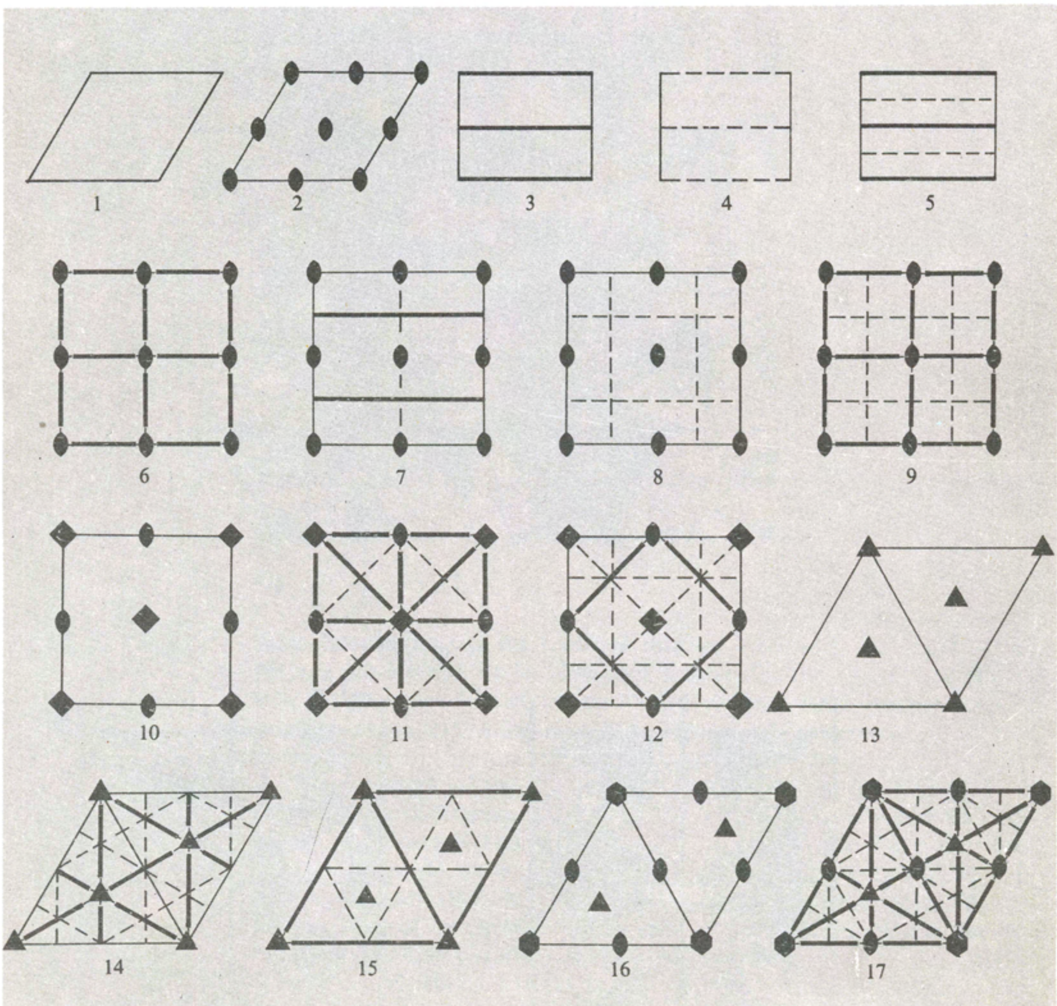


Fig. 35

Turning to the Egyptian design of Fig. 31, the main motif here may be the design within the confines of the shaded rectangle at the bottom of Fig. 34*a* (it makes up 1/8 of the surface area of the unit cell). The main motif is given separately in Fig. 36*a*. To construct a pattern we will make use of axes *BC* and *DE* in Figs. 34*c* and 36, and the two-fold rotation axis passing through point *A*. We will now fix *A* and *BC* and *DE* in the figure and obtain for the main motif (Fig. 36*a*) reflections relative to *BC* and *DE* and turns by 180° about *A* in any sequence and

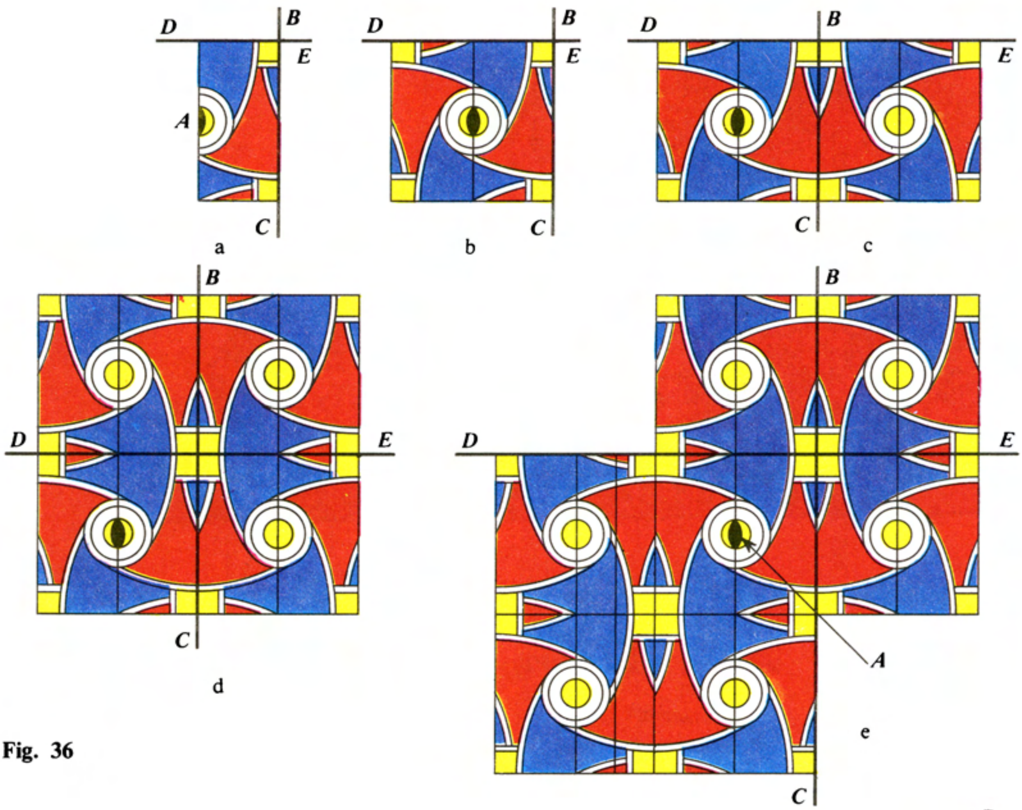


Fig. 36

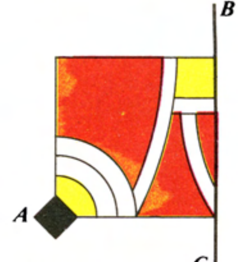


Fig. 37

indefinitely. A turn about A gives us Fig. 36*b*, a subsequent reflection relative to BC gives us Fig. 36*c*. Next we effect a reflection relative to DE (Fig. 36*d*) and another turn by 180° about A (Fig. 36*e*), yet another reflection relative to BC, and so on. As we go through the procedure, the pattern comes to life before our very eyes filling up the whole of the figure.

Let us now take the Egyptian design with a simplified colouring (Fig. 34*c*). For our main motif we can now have the double-shaded

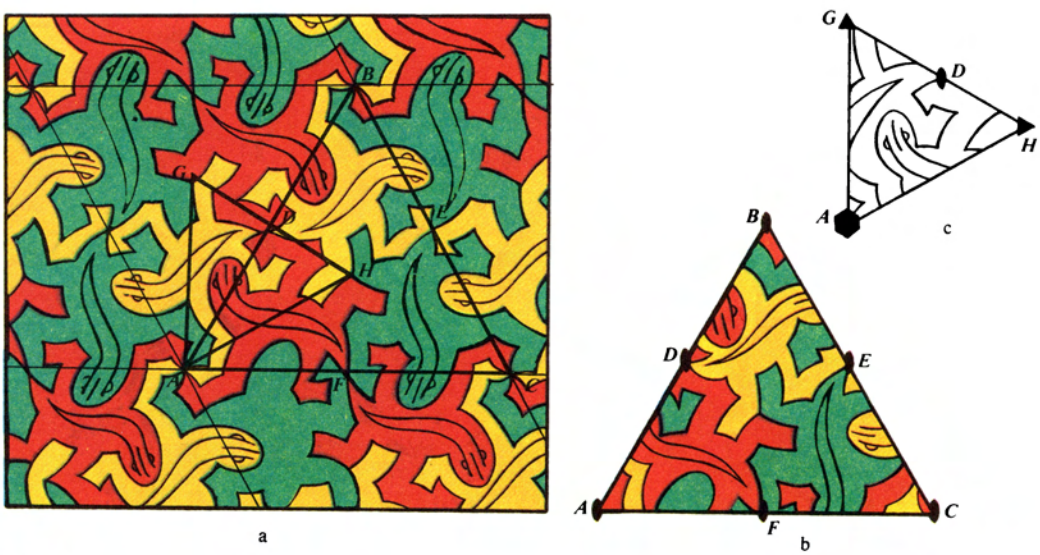


Fig. 38

square in Fig. 34*a*. We will construct our pattern using the four-fold axis and axis BC (Fig. 37). The reader may do this on his own.

The 'Lizards' Design

This design is quite interesting (Fig. 32). It is essentially a mosaic composed of identical lizards which are densely packed within the space of the pattern (without gaps or overlaps). The mosaic features not only translational but also *rotational* symmetry. The translational symmetry of the pattern is determined by the hexagonal lattice, and the rotational one by the rotation axes at A, B, C, D, E, F, G, H, etc. (Fig. 38*a*). The order of the rotational axes depends on the pattern's colouring. For a *three-colour* pattern (lizards of three different colours) all the rotation axes are two-fold (Fig. 38*b*). A *single-colour* pattern, apart from two-fold axes, has three- and six-fold rotation axes (Fig. 38*c*). The 'Lizards' have no mirror symmetry. The elements of symmetry for the single-colour design are given in # *16* in Fig. 35, and for the three-colour one in # *2*.

In constructing the single-colour 'Lizards', we can select as the main motif the triangle AGH in Fig. 38*c*. It is easily seen that the triangle includes parts of one lizard and its area is equal to that occupied by one lizard. The pattern can be constructed using the six-fold rotation axes (A) and the three-fold rotation axis (H). We will now turn the design in Fig. 38*c* through 60° about A, and then (after a complete turn) we will turn the resultant design about H through 120°.

In the case of the three-colour version the main motif is specified not by AGH but by triangle ABC that includes components of all three multicoloured lizards (Fig. 38*b*). The pattern can be obtained using the two-fold rotation axes that pass, for example, through points D, E, and F.

4 Regular Polyhedra

Inhabitants of even the most distant galaxy cannot play dice having a shape of a regular convex polyhedron unknown to us.

M. Gardner

Since time immemorial, when contemplating the picture of the Universe, man made active use of the *concept of symmetry*. Ancient Greeks preached a symmetrical universe simply because symmetry was beautiful to them. Proceeding from the concept of symmetry, they made some conjectures. So Pythagorus (the 5th century B.C.), who thought of the sphere as the most symmetrical and perfect shape, concluded that the Earth is spherical and moves over a sphere. He also assumed that the Earth revolves about some 'central fire', together with the six planets known at the time, the Moon, the Sun, and the stars.

The ancients liked symmetry and besides spheres they also used regular polyhedra. So they established the striking fact that there are *only five regular convex polyhedra*. First studied by the Pythagoreans, those five regular polyhedra were later described in more detail by Plato and so they came to be known in mathematics as *Platonic solids*.

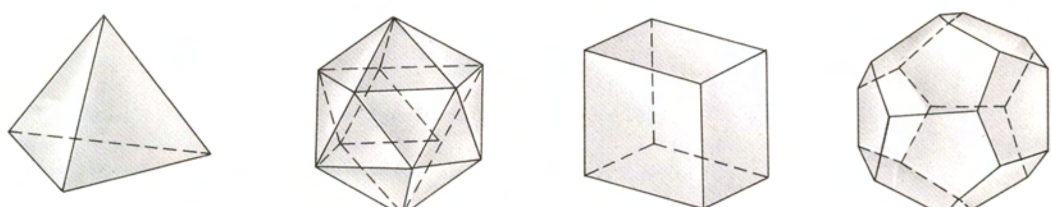
The Five Platonic Solids

A *regular polyhedron* is a volume figure with congruent faces having the shape of regular polygons, and with congruent dihedral angles. It turns out that there can be only five such figures (although there exist an infinite number of regular polygons). All the types of regular polyhedra are provided in Fig. 39: the *tetrahedron* (regular triangular pyramid), *octahedron*, *icosahedron*, *hexahedron* (cube), *dodecahedron*. The cube and octahedron are *mutual*: if in one of these polyhedra the centres of faces having a common edge are connected, the other polyhedron is obtained. Also mutual are the dodecahedron and icosahedron.

It can readily be shown why only five regular polyhedra are possible. Let us take the simplest face – the *equilateral triangle*. A polyhedral angle can be obtained by putting together three, four or five equilateral triangles, that is, in three ways. (If there are six triangles, the angles at the common vertex sum up to 360°.) With *squares* a polyhedral angle can only be formed in one way – with three squares. With *pentagons* too, a polyhedral angle can be obtained in one way – with three pentagons. Clearly, no regular polyhedral angles can be formed out of n-hedra with $n \geqslant 6$. Consequently, only five types of regular polyhedra can exist: three polyhedra with triangular faces (tetrahedron, octahedron, and icosahedron), one with square faces (cube), and one with pentagon faces (dodecahedron).

The Symmetry of the Regular Polyhedra

The elements of *symmetry of the tetrahedron* are illustrated in Fig. 40. It has four three-fold rotation axes and three two-fold axes. Each of the three-fold axes passes through a vertex and the centre of the opposite face (for example, axis *AE* in the figure). Each two-fold axis passes through the middles of opposite edges (for example, axis *FG*). Through each two-fold axis pass two planes of symmetry (through an axis and one of the edges that intersect with a given axis); in the figure planes *AGC* and *DFB* pass through axis *FG*. The tetrahedron thus has six planes of symmetry. In addition, the tetrahedron has a mirror-rotational symmetry: each two-fold rotation axis also doubles as a four-fold mirror-rotation axis.



Tetrahedron Icosahedron Cube Dodecahedron Octahedron

Fig. 39

The *symmetry of the cube* was discussed partially in Chapter 2. Recall that the cube has 13 rotation axes of symmetry: three four-fold axes, four three-fold axes, and six two-fold axes. Interestingly, the *octahedron* has the same elements of *symmetry* as the cube. The octahedron has three four-fold rotation axes (they pass through opposite vertices, like axis *AB* in Fig. 41), four three-fold axes (they pass through the centres of opposite faces, like axis *CD*), and six two-fold axes (they pass through centres of opposite, mutually parallel edges, like axis *EF*). Both the cube and the octahedron have nine planes of symmetry (find them on your own). Lastly, each three-fold rotation axis in the cube and the octahedron is at the same time a six-fold mirror-rotation axis.

It has already been noted above that cube and octahedron are *mutual* polyhedra. That is why they have the same elements of symmetry. The mutuality of the dodecahedron and the icosahedron also implies that the two have the same symmetry.

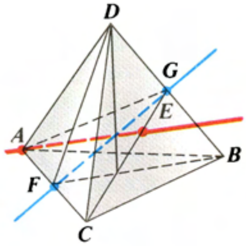


Fig. 40

The *symmetry of the dodecahedron* is illustrated in Fig. 42. Axis *AB*, which passes through the centres of the opposite faces, is one of the six five-fold rotation axes; axis *CD*, which passes through the opposite vertices, is one of the ten three-fold axes; axis *EF*, which passes through the centres of the opposite mutually parallel edges, is one of the 15 two-fold rotation axes. The same rotation axes are present in the *icosahedron*; only in it five-fold axes pass through opposite vertices, not through the centres of opposite faces, whereas three-fold axes pass through the face centres.

The Uses of the Platonic Solids to Explain Some Fundamental Problems

The notion of symmetry has often been used as the framework of thought for hypotheses and theories of scholars of centuries past, who put much stock in the mathematical harmony of the creation of the world and regarded that harmony as a sign of divine will. They viewed the Platonic solids as a fact of fundamental importance, directly related to the structure of matter and the Universe.

So the Pythagoreans, and later Plato, believed that matter consists of four principal elements–*fire*, *earth*, *air*, and *water*. According to their thinking, the atoms of principal elements must have the shape of various Platonic solids: fire atoms must be tetrahedra; earth atoms, cubes; air atoms, octahedra; and water atoms, icosahedra.

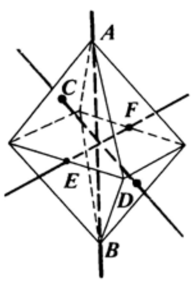


Fig. 41

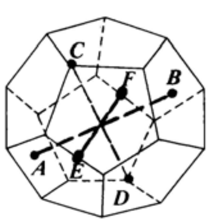


Fig. 42

The concept of symmetry coded in the five Platonic solids enthralled the famous German astronomer Johannes Kepler (1571-1630), who undertook to explain why there are just six planets in the Solar system (in Kepler's days, too, only six planets were known), and why the radii of their 'spheres' (orbits) are in the ratio 8 : 15 : 20 : 30 : 115 : 195 (according to Kepler's results). Kepler inscribed a cube into Saturn's sphere. Next into that cube he inscribed another sphere, that of Jupiter. Into Jupiter's sphere he inscribed a tetrahedron, and into the tetrahedron the Martian sphere, and into the latter a dodecahedron. Finally, he inscribed the Earth's sphere. Then followed in succession an icosahedron, inscribed into the Earth's sphere, Venus's sphere, an octahedron inscribed into the latter, and lastly Mercury's sphere. It is easily seen that Kepler's scheme employs all five Platonic solids. Part of the scheme is depicted in Fig. 43. Kepler calculated the radii of the planetary spheres in accordance with his scheme to find that the ratio of these radii were in good agreement with the observations. This striking coincidence made Kepler believe in his underlying assumption. He thought that he had succeeded in explaining the structure of the entire Solar system on the basis of a geometrical system using spheres and the five Platonic solids, thus directly relating the existence of six planets to the existence of the five Platonic solids. Kepler wrote: "The great joy I experienced from that discovery cannot be put into words. I did not regret any of the time spent and felt no fatigue. I was not afraid of cumbersome calculations while seeking to find whether my hypothesis was in agreement with Copernicus's theory of orbits, or whether my joy was to vanish into smoke."

Kepler's enthusiasm turned out to be premature. The coincidence he hit upon was accidental and, as was shown by later observations, quite rough. And to top it off, there are actually nine planets, not six, in the Solar system.

On the Role of Symmetry in the Cognition of Nature

The two just-discussed examples of unsuccessful applications of the Platonic solids to explain fundamental issues of nature suggest that

symmetry alone is not sufficient. Yet the uses of symmetry to ponder the world around us are significant.

The Platonic solids simply furnish an example of how symmetry may drastically limit the variety of structures possible in nature. We will now elaborate this important idea. Suppose that in some distant galaxy, live highly developed creatures who, among other things, are fond of games. We may be totally ignorant of their tastes, structure of their bodies, and their mentality. We know, however, that their dice have any of the five

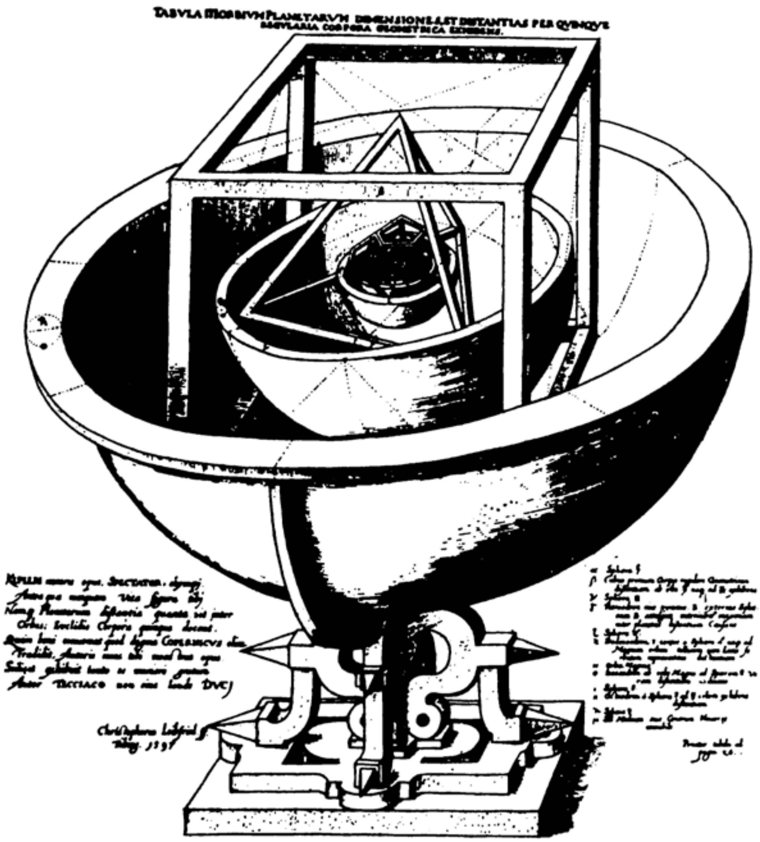


Fig. 43

shapes – tetrahedron, cube, octahedron, dodecahedron, and icosahedron. Any other shape of a dice is impossible in principle, since the requirement that any face must have equal probability of appearance dictates the use of a regular polyhedron, and there are only five regular polyhedra.

A measure of order introduced by symmetry shows up, above all, in limitation of the variety of possible structures and reduction of the number of possible versions.

Later in the book we will discuss some limitations imposed by symmetry on the diversity of structures of molecules and crystals. The modern American popularizer of science Martin Gardner writes: "Perhaps some day physicists will discover some mathematical constraints to be met by the number of elementary particles and basic laws of nature."

5 Symmetry in Nature

We like to look at symmetrical things in nature, such as perfectly symmetrical spheres like planets and the Sun, or symmetrical crystals like snowflakes, or flowers which are nearly symmetrical.

R. Feynman

From the Concept of Symmetry to the Real Picture of a Symmetrical World

We have seen that the concept of symmetry has often been used by scientists through the ages as a guiding star in their speculations. Recall the Pythagoreans who concluded that the Earth is spherical and moves over a sphere. Pythagoras's ideas were used by the great Polish astronomer Copernicus when he was working on his theory of the Solar system. According to Copernicus, celestial bodies are spherical because the sphere is 'a perfect, comprehensive shape, having no corners, the most capacious'. He wrote: "All bodies tend to assume that form; this can be noticed in water droplets and other liquid bodies." Here Copernicus meant free falling drops, which are known to take on a nearly spherical shape. In actual fact, he anticipated the deep analogy between a water drop falling under gravity and the Earth falling (or rather orbiting) in the gravitational field of the Sun.

On the other hand, scholars of earlier times were inclined to exaggerate a bit the role of symmetry in the picture of the world. They sometimes forced their admiration for symmetry on nature by artificially squeezing nature into symmetrical models and schemes. Recall Kepler's scheme based on the five regular polyhedra.

The modern picture of the world, with its rigorous scientific justification, differs markedly from earlier models. It excludes the existence of some 'centre of the world' (just as some magic power of the Platonic solids) and treats the Universe in terms of the *unity of symmetry and asymmetry*. Observing the chaotic mass of stars in the skies, we understand that beyond the seeming chaos are quite symmetrical spiral structures of galaxies, and in them symmetrical structures of planetary systems. This symmetry is illustrated in Fig. 44 showing the Galaxy and a magnified and simplified scheme of the Solar system.

The nine planets move around the Sun in their elliptical orbits which are nearly circular. The planes of the orbit (save for Pluto) within high accuracy coincide with the Earth's orbital plane, the so-called *ecliptic*. For example, Mars's orbit forms an angle of 2° with the ecliptic. Also coinciding with the ecliptic are the planes of all the 13 satellites of the planets, including the Moon.

Even more than in the picture of the universe, symmetry manifests itself in an infinite variety of structures and phenomena of the inorganic world and animate nature.

Symmetry in Inanimate Nature. Crystals

When we look at a heap of stones on a hill, the irregular line of mountains on the horizon, meandering lines of river banks or lakeshores, the shapes of clouds, we may think that symmetry in the inorganic world is rather rare. At the same time, it is widely believed that symmetry and strict order are hostile to living things. It is no wonder that the lifeless castle of the Snow Queen in the fairy tale by Hans Christian Andersen is often pictured as a highly symmetrical structure shining with polished

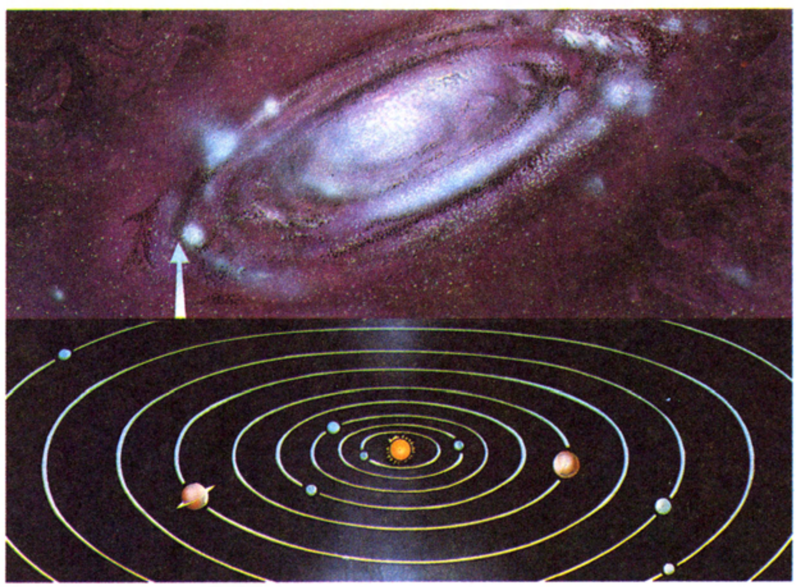


Fig. 44

mirror faces of regular shape. Who is right then? Those who view inanimate nature as a realm of disorder, or, on the contrary, those who see in it the predominance of order and symmetry?

Strictly speaking, both schools are wrong. To be sure, such natural factors as wind, water, and sunlight affect the terrestrial surface in a highly random and disorderly manner. However, sand dunes, pebbles on the seashore, the crater of an extinct volcano are as a rule regular in shape. Of course, a heap of stones is rather disorderly, but each stone is a huge colony of *crystals*, which are utterly symmetrical structures of atoms and molecules. *It is crystals that make the world of inanimate nature so charmingly symmetrical.*

Inhabitants of cold climates admire *snowflakes*. A snowflake is a tiny crystal of frozen water. Snowflakes are of various shapes, but each has a six-fold rotational symmetry and also a mirror symmetry (Fig. 45).

All solids consist of *crystals*. Individual crystals are generally tiny (less

than a grain of sand), but in some cases they grow to considerable sizes, and then they appear before us in all their geometrical beauty. Some naturally grown crystals are given in Fig. 46. It can be seen from the figure that crystals are polyhedra of fairly regular shapes having plane faces and straight edges. The figure shows *topaz* (aluminium fluosilicate), *beryl* (beryllium aluminium silicate), *smoky quartz* (silicon dioxide).

The beryl crystal shown in the figure is *heliodor*, one of the crystalline varieties of that compound. Other varieties are *aquamarine* (blue-green),

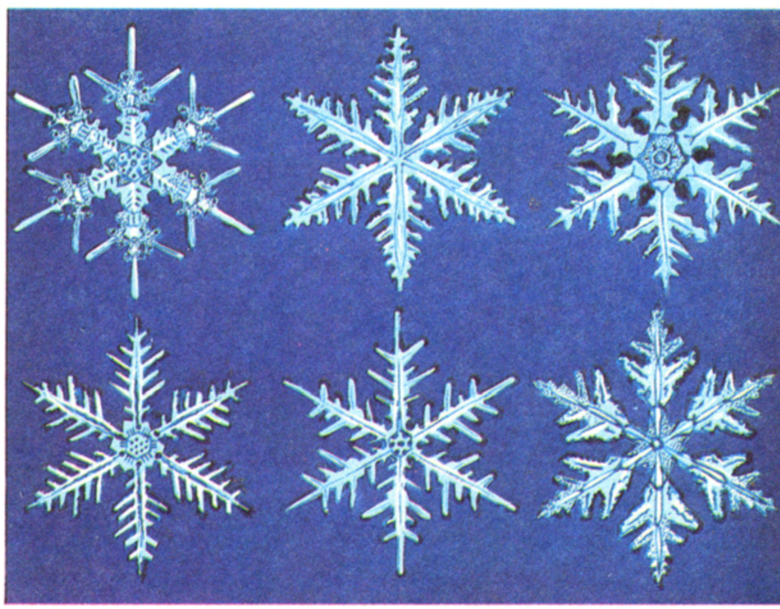
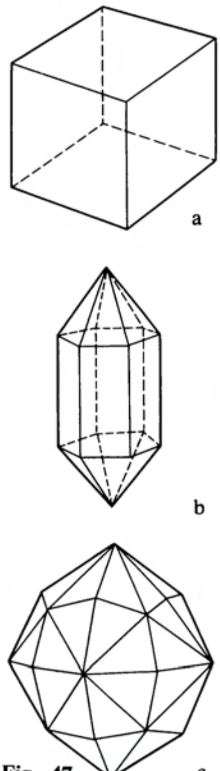


Fig. 45

Fig. 46 TOPAZ SMOKY QUARTZ BERYL



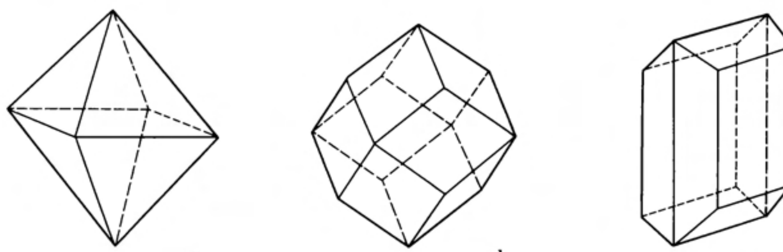
a

b

Fig. 47 c

emerald (green), and *vorobyevite* (pale red). The colour is conditioned by impurities. So, the yellow of heliodor is due to Fe^{3+} impurities. Quartz comes in a variety of forms. The clearest and transparent variety of quartz is *rock crystal*, the second clearest is *smoky quartz*, shown in the figure. There are also *violet amethyst*, *red sardonyx*, *black onyx*, and *grey chalcedony*. Quartz is also grindstone, flint, and common sand.

The symmetry of crystals is clearly seen in Fig. 47: (a) *common salt*, (b) *quartz*, and (c) *aragonite*. The latter is one of the naturally occurring



a b c

Fig. 48

varieties of *calcite* ($CaCO_3$). Figure 48 represents three crystalline forms of *diamond*: (a) octahedron, (b) rhombic dodecahedron, and (c) hexagonal octahedron.

The *outer symmetry* of a crystal comes from its inner symmetry, that is, its ordered arrangement of atoms (molecules) in space. In other words, the symmetry of a crystal is related to the existence of the space lattice of atoms – the so-called *crystalline lattice*.

Symmetry in the World of Plants

In his book *The Ambidextrous World*, M. Gardner writes: "On the earth life started out with spherical symmetry, then branched off in two major directions: the plant world with symmetry similar to that of a cone, and the animal world with bilateral symmetry."*

The *symmetry of a cone*, which is characteristic of plants, is seen essentially in any tree (Fig. 49). The root of a tree absorbs water and nutrients from the soil, that is, from *below*, whereas all the other vital functions occur in the canopy, that is, *above* the ground. Therefore, the directions 'upwards' and 'downwards' for the tree are significantly different. At the same time the direction in a plane perpendicular to a vertical are essentially indistinguishable: from every quarter the tree receives air, light, and water in equal measure. As a result, we have a vertical rotation axis (the cone's axis) and vertical planes of symmetry. Note that the vertical orientation of the cone's axis, which characterizes

Fig. 49

* The term 'bilateral symmetry' is widely used in biology. It means mirror symmetry.

the symmetry of a tree, is determined by the direction of *gravity**. That is why the general orientation of the stem of a tree is normally independent of the slope of the ground or the Sun's altitude in a given latitude.

True, trees can be encountered now and then, such that their stems are not just nonvertical, but bent in a snake-like manner, and their canopy may be lop-sided. It would seem that any symmetry here is out of the question. And still the concept of a cone at all times correctly reflects the nature of the symmetry of the tree, its essence. After all, for any tree we

Fig. 50

can indicate the *base* and the *top*, and at the same time for the tree the notions of left and right, back and front are invalid.

Remarkable symmetry is inherent in leaves, branches, flowers, and fruit. Figure 50 gives examples of mirror symmetry, which is characteristic of leaves, although is also encountered in flowers, which are generally described by rotational symmetry. Figure 51*a* depicts a flower of *St. John's wort* (*Hypericum*), which has a five-fold rotation axis and no mirror symmetry. In flowers, rotational symmetry is often accompanied by mirror symmetry (Fig. 51*b*). An *acacia* leaf, shown in Fig. 52*a*, has

---

* Biological experiments on board the Soviet orbital station *Salyut-6* have shown that under weightlessness, the spatial orientation of sprouts, leaves, and roots of wheat and peas is violated.

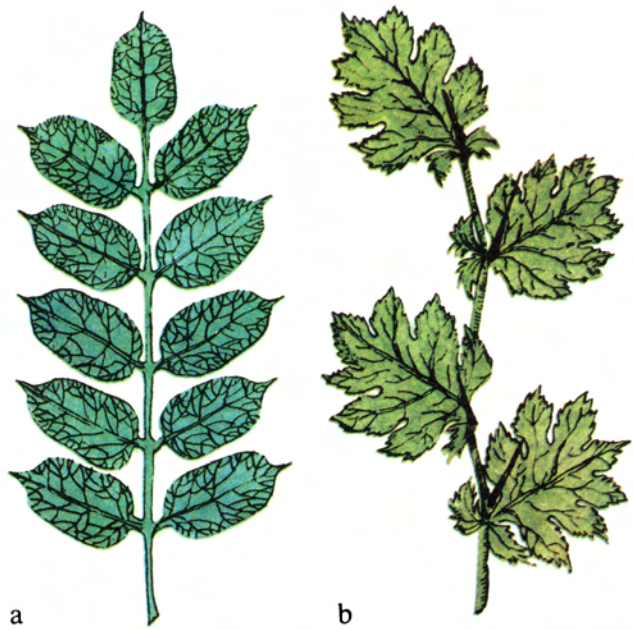
both mirror and translational symmetries. And a *hawthorn* branch (Fig. 52*b*) can be seen to have a glide axis of symmetry.

Figure 53 shows a wild flower known as *silverweed* (*Potentilla anserina*). The flower has a five-fold rotation axis and five planes of symmetry. The high degree of orderliness in the arrangement of individual leaves on stems imparts to the figure some likeness to borders discussed above.

a

b

Fig. 51



a b

Fig. 52

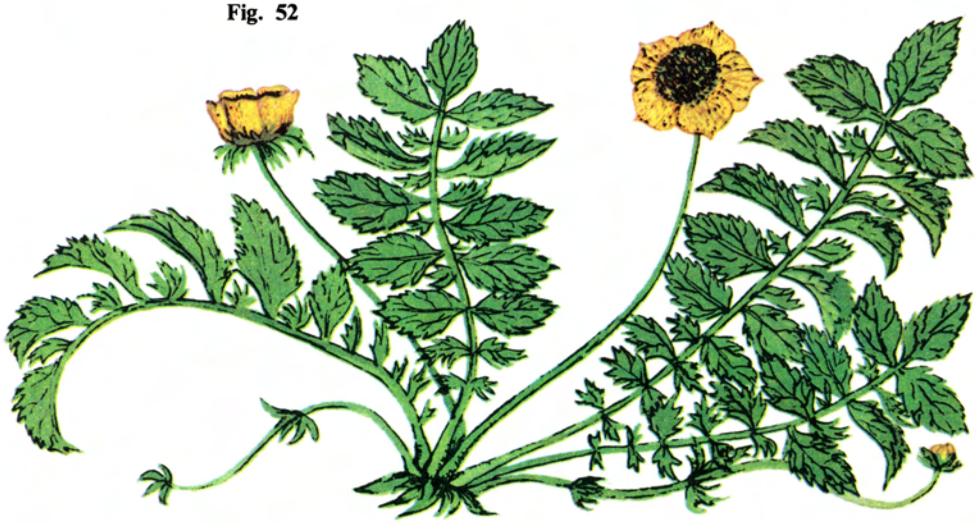


Fig. 53

In the rich world of flowers rotation axes of symmetry of various orders occur. But the most widespread is the five-fold rotational symmetry. This symmetry is to be found in many wild flowers (*bluebell, forget-me-not, geranium, stellaria, pink, St. John's wort, silverweed,* etc.), in fruit tree flowers (*cherry, apple, pear, mandarin,* etc.), in flowers of fruit and berry plants (*strawberry, blackberry, raspberry, guelder rose, bird-cherry, rowan-tree, hawthorn, dog-rose,* etc.), and in some garden flowers (*nastur-*

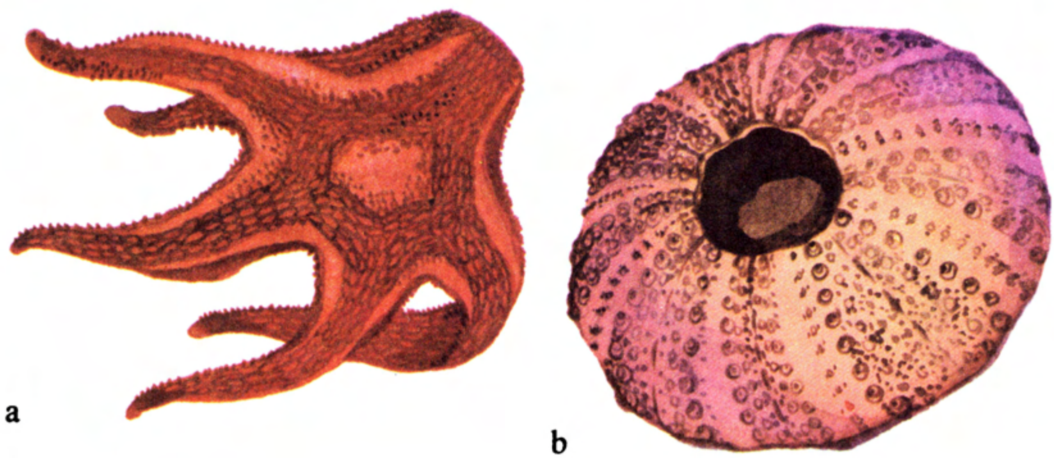


Fig. 54

tium, phlox, etc.). It is sometimes argued that plants' 'love', which is known to be impossible in principle in periodic structures, can be explained as a safeguard of the plant's individuality. Academician N. Belov maintains "that the five-fold axis is a sort of tool in the struggle for existence, an insurance against petrification, the first stage of which would be 'catch' by a lattice."

Symmetry in the World of Animals

The five-fold rotational symmetry also occurs in the world of animals. Examples are the *starfish* and the *urchin* (Fig. 54). Unlike the world of plants, however, rotational symmetry is rare in the world of animals. As a matter of fact, we may find it only in some denizens of the sea.

Insects, fishes, birds and other animals generally exhibit a difference between forward and backward directions, which is incompatible with rotational symmetry. The Push-Pull invented in a famous Russian fairy tale (Fig. 55) is a striking animal in that its front and rear are absolutely symmetrical. The *direction of motion* is an essentially distinguishable direction, about which no animal is symmetrical. In that direction an animal moves for its food and escapes from danger.

The symmetry of living creatures is also dictated by another direction – the *direction of gravity*. Both directions are significant, since they define the *plane of symmetry* of a creature (Fig. 56). *Bilateral* (*mirror*) *symmetry is characteristic of nearly all members of the animal kingdom.*

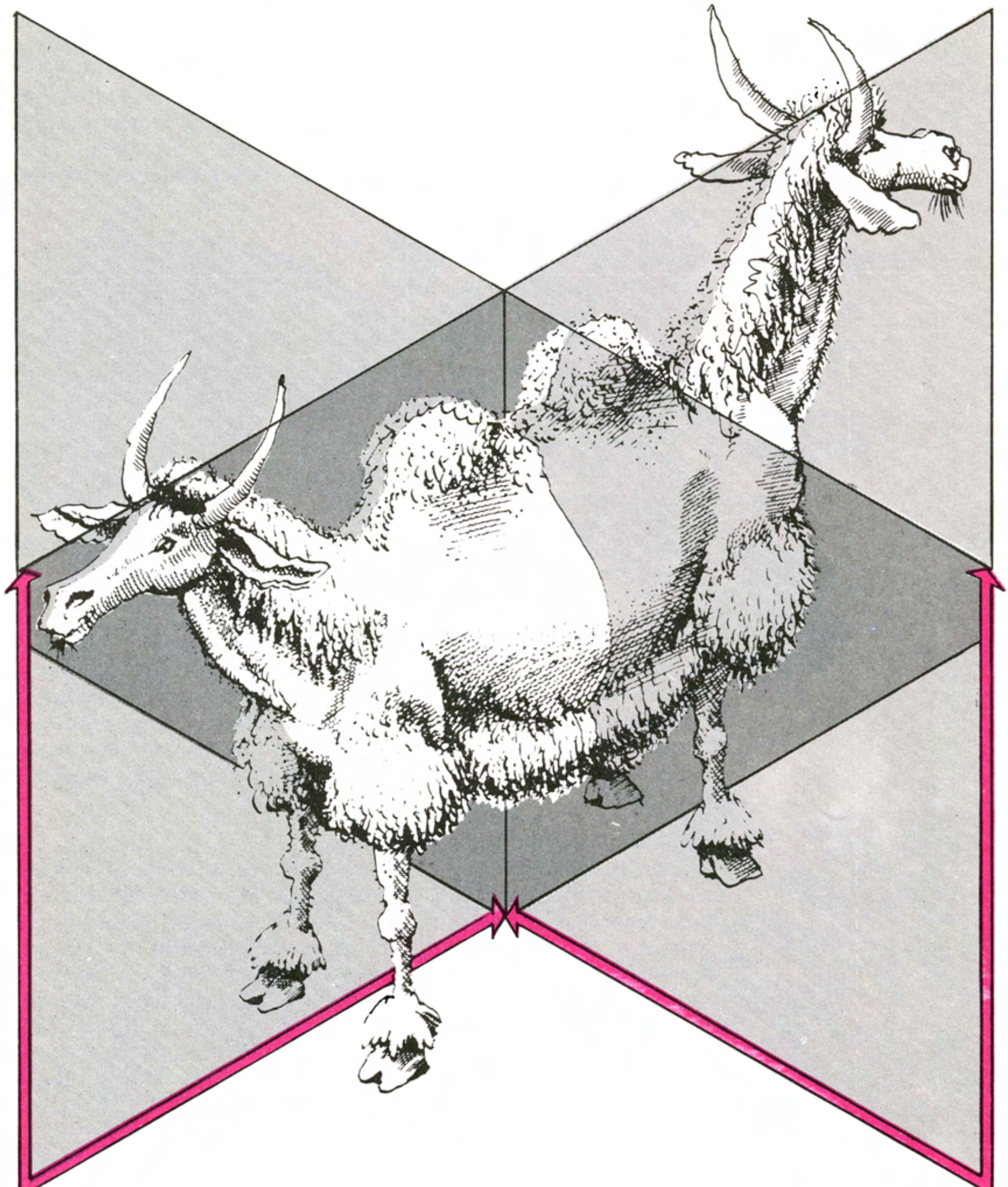


Fig. 55

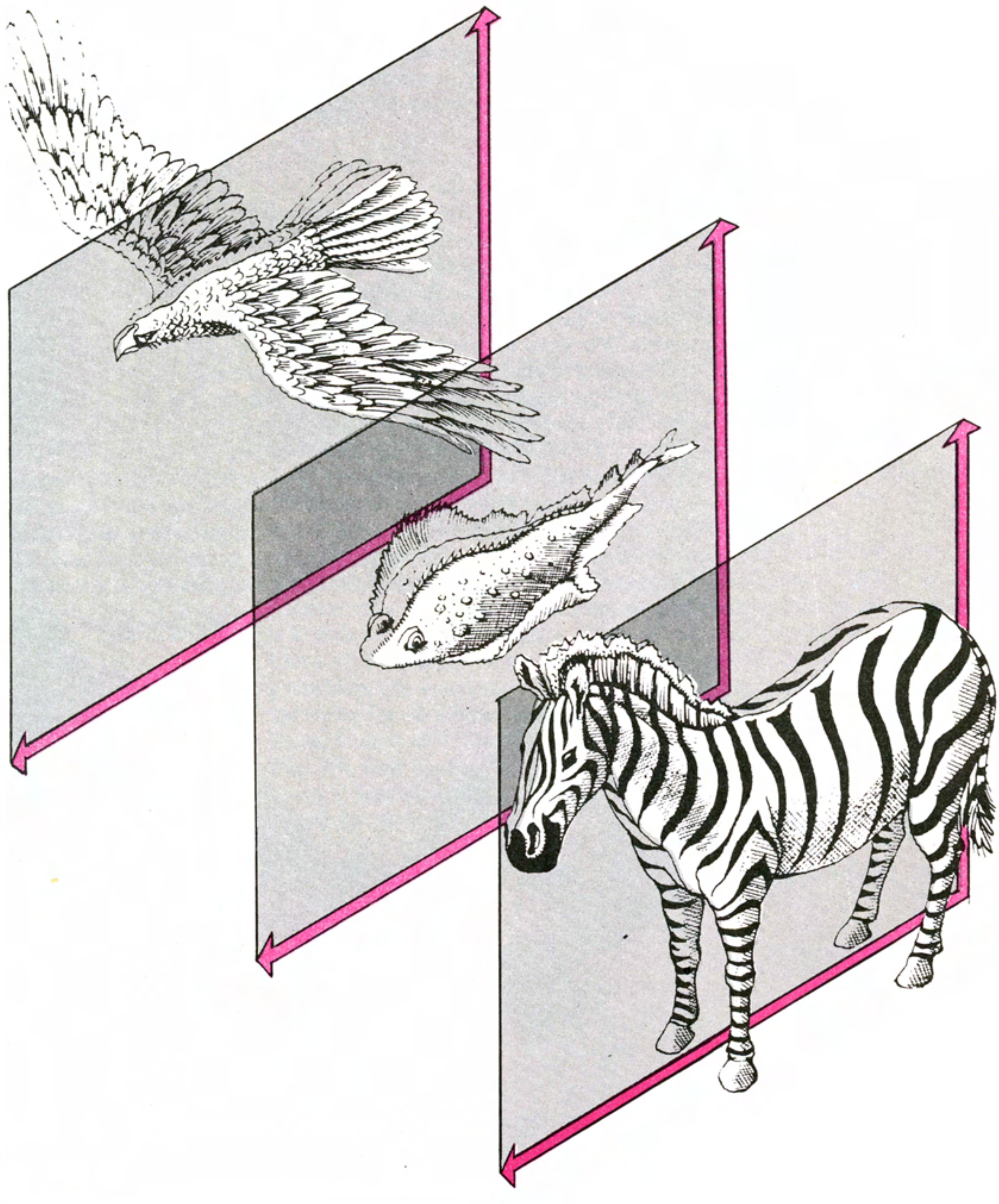


Fig. 56

This symmetry is especially apparent in a *butterfly* (Fig. 57). The symmetry of left and right is present here with nearly mathematical accuracy.

It can be said that any animal consists of two *enantiomorphs* – its left and right halves. Also enantiomorphs are paired organs, one of which is in the right and the other in the left half of the body, such as ears, eyes, horns, and so on.

Inhabitants of Other Worlds

Many works of science fiction discuss the possible appearances of visitors from other planets. Some authors believe that extraterrestrials may differ markedly in their appearance from 'earthlings'; others, on the contrary, believe that intelligent creatures throughout the entire Universe must be very much alike. The question concerns us only in the context of symmetry. Whatever the extraterrestrial looks like, his appearance must exhibit bilateral symmetry, because on any planet a living creature must have a distinguishable direction of motion and on any planet there is gravity. The extraterrestrial may be like a dragon from some fairy tale, but not like a Push-Pull, by no means. He cannot be left-eyed or right-eared. He must have an equal number of limbs on either side. Symmetry requirements reduce drastically the number of possible versions of the extraterrestrial's appearances. And although we cannot say with certainty what that appearance *must be*, we can say what it *cannot be*. Recall the idea expressed in Chapter 4: *symmetry limits the diversity of structures possible in nature.*

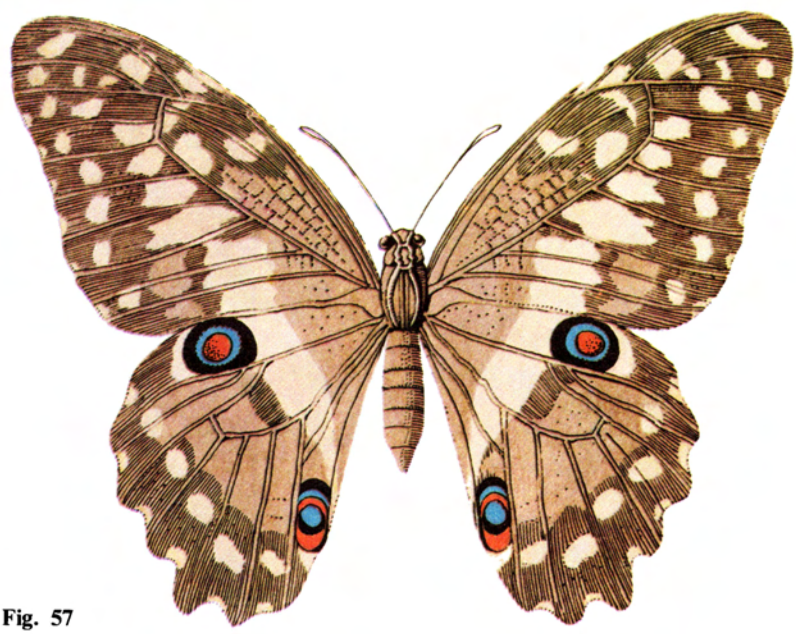


Fig. 57

6 Order in the World of Atoms

The crystal is characterized by its internal structure, arrangement of atoms, and not by its outward appearance. These atoms combine to produce a sort of huge molecules, or rather an ordered space lattice.

H. Lindner

Molecules

Chapter 5 was concerned with symmetry in nature as seen by the unaided eye. Symmetry is also found at the *atomic level*. It manifests itself in microscopic, geometrically ordered atomic structures of molecules and crystals.

Figure 58 presents schematically two simple molecules: (a) *carbon dioxide* (CO_2), and (b) *steam* (H_2O). Both molecules have a plane of symmetry (the vertical line in the figure). The mirror symmetry here comes from the fact that paired identical atoms (oxygen atoms in CO_2 or hydrogen atoms in H_2O) are bound to the third atom in a similar way. Interchanging the paired atoms will change nothing – this will only amount to mirror reflections.

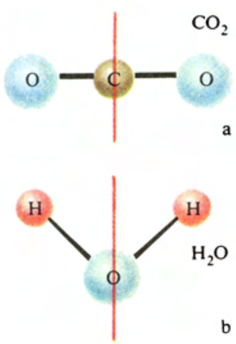


Fig. 58

In the *methane* molecule (CH_4), the carbon atom C is bound to the four identical hydrogen atoms H. The four C—H bonds being identical predetermines the spatial structure of the molecule in the shape of a *tetrahedron*, with hydrogen atoms being at the corners and a carbon atom at the centre (Fig. 59). The symmetry of the molecule CH_4 is essentially the symmetry of the tetrahedron discussed in Chapter 4. Its elements are six planes of symmetry, each of which passes through the atom C and two atoms H (for example, planes LMQ and LKR in Fig. 60), four three-fold rotation axes, each of which passes through the atom C and one of the atoms H (for example, axis KO in the figure). There are three two-fold rotation axes (for example, axes PQ and SR). Notice the difference between the *spatial structure* of the methane molecule given in Fig. 59 and the *structural formula* of the molecule

$$\begin{array}{ccc} & \mathrm{H} & \\ & | & \\ \mathrm{H} - & \mathrm{C} & - \mathrm{H} \\ & | & \\ & \mathrm{H} & \end{array}$$

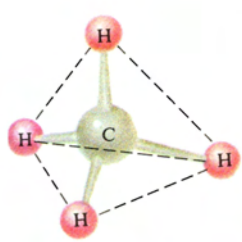


Fig. 59

normally given in chemistry texts. Now suppose that one of the hydrogen atoms in the molecule (for example, the atom at K in the tetrahedron shown in Fig. 60) is replaced by a radical OH. In that case we obtain the molecule of *methyl alcohol* (CH_3OH). As compared with the methane molecule, this molecule exhibits *lower* symmetry (even supposing that the molecule retains its tetrahedral shape). It is easily seen that out of the six

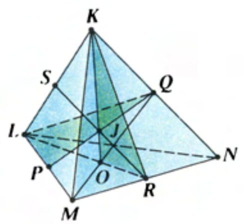


Fig. 60

planes of symmetry only three remain – those passing through C, OH, H (planes *LKR*, *KPN*, *KMT* in Fig. 61); out of the four three-fold rotation axes, only one remains (*KO* in the figure). There are no two-fold rotation axes now.

Notice that the chemical formula of methyl alcohol is now written CH_3OH, not CH_4O. This is no mere chance. The form CH_4O would mean that all four H atoms in the molecule are physically equivalent, which is not the case here: only three H atoms are equivalent, whereas the fourth stands alone, as it enters the OH radical.

The Puzzle of the Benzene Ring

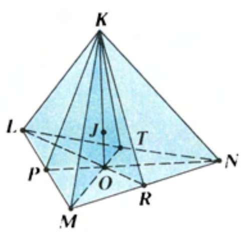


Fig. 61

The *benzene* molecule consists of six carbon and six hydrogen atoms (C_6H_6). The carbon atoms are arranged in one plane to form a *regular hexagon* (the so-called *benzene ring*). It is well known that carbon is tetravalent, that is, a carbon atom provides four electrons which can realize *four* covalent bonds with other atoms. One of them is the bond between a carbon atom and a hydrogen atom, the other three binding a given atom to neighbouring carbon atoms in the benzene ring. The structural formula of the benzene molecule is sometimes presented as shown in Fig. 62, where some pairs of carbon atoms are bound by *single* and others by *double* bonds.

At first sight everything about the structural formula given in Fig. 62 is OK. But this is only true at first sight. The fact is that the presence of different bonds (single and double) must violate the regular shape of the benzene ring, since stronger (double) bonds correspond to smaller atomic spacings. At the same time, X-ray studies show that all the sides of the carbon hexagon in the benzene molecule are equal. The experimentally found symmetry of the benzene ring (the symmetry of the regular hexagon) suggests that all the C—C bonds in the ring are equal. What is the nature of these bonds? These cannot be single covalent bonds, otherwise one bond in each carbon atom would be free. But they cannot be double bonds either, since for this to be the case each carbon atom lacks one valence.

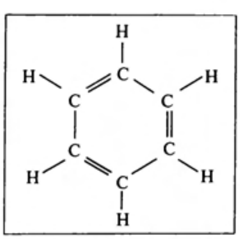


Fig. 62

The enigma of the benzene ring turned out to be exceedingly interesting. One of the valence electrons of each carbon atom participates in the formation of a bond of this atom with five atoms of the ring at once, and not with one of the neighbouring atoms. This implies that the electron is collectivized not by a pair of atoms (which is common for covalent bonds), but by the *entire molecule* (or rather the entire benzene ring). In other words, the benzene molecule has six electrons not pinned up by localized bonds between atoms, but capable of freely moving around the entire benzene ring. This is generally represented by the structural formula shown in Fig. 63, where the solid lines denote, as usual, the localized bonds (each bond due to the collectivization of a pair of electrons by a pair of appropriate atoms), and dash lines denote nonlocalized bonds due to the collectivization of six electrons by the benzene ring.

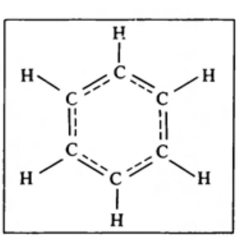


Fig. 63

The Crystal Lattice

It would seem that *diamond* and *graphite* have nothing in common. The diamond is unusually hard, transparent, and is a dielectric. Processed stones are used in jewelry. Graphite, on the other hand, is soft, laminar, opaque, electroconductive. In a word, a far cry from a gem. At the same time, however, both diamond and graphite are *carbon* in its pure form. The different behaviour of diamond and graphite is only explained by their different crystalline structure, or *different crystalline lattices.* This is a graphic example of the important role played by the crystalline lattice in determining the properties of a solid.

The crystal lattice is a natural three-dimensional pattern. As with plane patterns, it is dominated by some form of translational symmetry. It has been noted in Chapter 2 that there exist 14 types of spatial lattices that differ in their translational symmetry (14 types of *Bravais lattices*). They form seven *crystallographic systems*:

cubic system: $a = b = c,\ \alpha = \beta = \gamma = 90°$;
tetragonal system: $a = b \neq c,\ \alpha = \beta = \gamma = 90°$;
hexagonal system: $a = b \neq c,\ \alpha = \beta = 90°,\ \gamma = 120°$;
trigonal system: $a = b = c,\ \alpha = \beta = \gamma \neq 90°$;
rhombic system: $a \neq b \neq c,\ \alpha = \beta = \gamma = 90°$;
monocline system: $a \neq b \neq c,\ \gamma \neq \alpha = \beta = 90°$;
tricline system: $a \neq b \neq c,\ \alpha \neq \beta \neq \gamma$.

Here a, b, and c are lengths of the edges of a unit cell; and α, β, and γ are angles between the edges (see Fig. 25).

The Face-Centred Cubic Lattice

Suppose that we have many balls of the same diameter. We will pack them densely on a plane (the brown balls in Fig. 64). Over the *first* layer a *second* one will be placed (the red balls). It is easily seen that the second layer is packed as densely as the first one. Next we will lay a *third* layer. Here we have two versions: (a) the centres of the balls in the third layer come exactly over the centres of the balls in the first layer; (b) the centres of the balls of the third layer are displaced horizontally relative to the balls in the first layer. Let us take the second version (the blue balls). The resultant dense multilayer packing corresponds to the *face-centred cubic* lattice (f. c. c.). In other words, the centres of the balls here (second version) form an f. c. c. lattice.

Figure 65 shows the *cubic* unit cell of the f. c. c. lattice. The sites here are the vertices of the cube and the centres of all its faces. We can readily discern in the figure the planes of the angles corresponding to the ball layers in Fig. 64. Let plane DEF correspond to the first layer (the brown balls). Then plane ABC will correspond to the second layer (the red balls). Site K will now belong to the third layer.

Each cell thus includes four sites (for example, sites D, P, M, and S in Fig. 65); the remaining sites in the figure must be assigned to neighbouring cells, and so the cell in question is a *four-site* cell.

The f. c. c. cell, just like any Bravais lattice, can also be defined using a *one-site* unit cell. Shown in Fig. 66 (in red) is a one-cell rhombohedron-

shaped cell. Crystallographers prefer using not one- but four-site cells, since it reflects in the most complete manner the elements of symmetry possessed by the f. c. c. cell.

Face-centred cubic lattices occur fairly often. This sort of lattice is to be found, for example, in *aluminium, gold, copper, nickel, platinum, silver*, and *lead*. The lattice of *common salt* (NaCl) actually consists of two geometrically identical interlocked f. c. c. lattices, one made up of Na^+ ions and the other of Cl^- ions (Fig. 67).

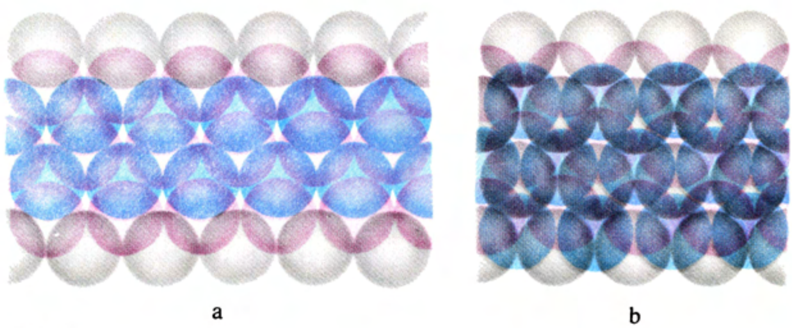


Fig. 64

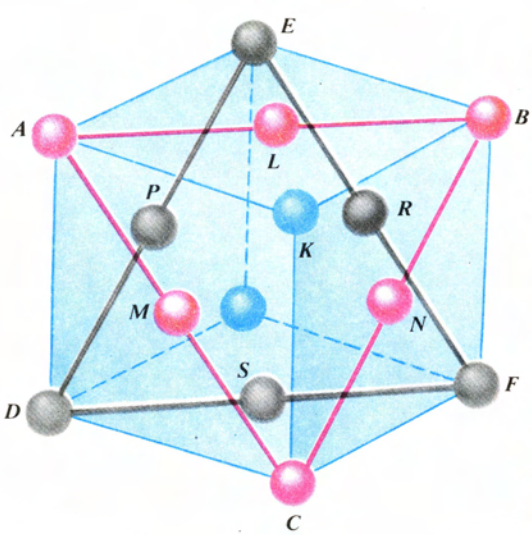


Fig. 65

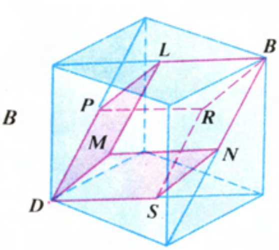


Fig. 66

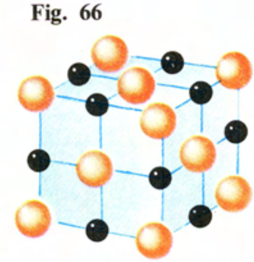


Fig. 67

Polymorphism

It has earlier been noted that the difference in the properties of diamond and graphite is determined by the difference in the crystal lattices of these two forms of carbon. As is seen in Fig. 68, the *diamond lattice* is formed by two identical interlocked f.c.c. lattices, one of which is displaced relative to the other by a quarter of the edge of the f.c.c. cell along all three coordinate axes (the white circles in the figure show sites of one of the f.c.c. lattices, and the filled circle shows the site of the other f.c.c.

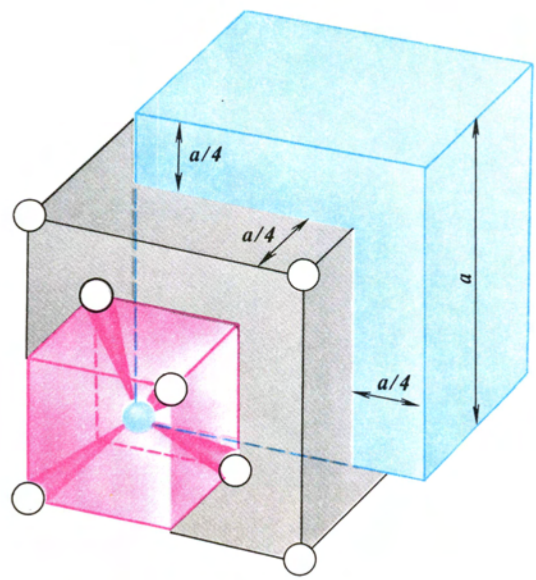


Fig. 68

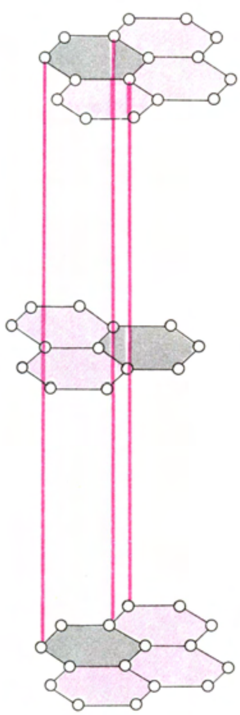


Fig. 69

cell). Each carbon atom in the diamond lattice is the centre of a *tetrahedron*, whose vertices are the four nearest neighbours of a given atom; this is seen especially clearly in the cell shown in red in the figure.

Note that the crystal lattices of *germanium, silicon* and *grey tin* have the diamond lattices.

The *graphite lattice* is given in Fig. 69. It is distinctly laminar in structure, with each layer being characterized by six-fold rotational symmetry. The bonds between the atoms from different layers are much weaker than within the same layer.

Diamond and graphite are good examples of two different crystalline modifications of a chemical element (or compound). This phenomenon is known as *polymorphism*. Under certain conditions a substance may change from one crystal modification to the other, and the changes are called *polymorphic transformations*. If, for instance, we heat graphite up to

2000-2500 K under a pressure of up to 10^{10} Pa, the crystal lattice will transform with the result that graphite will turn into diamond. In this way artificial diamonds are produced.

The Crystal Lattice and the External Appearance of a Crystal

The symmetry of the external shape of a crystal is conditioned by the symmetry of the crystal lattice. Ideally plane crystal faces are the planes

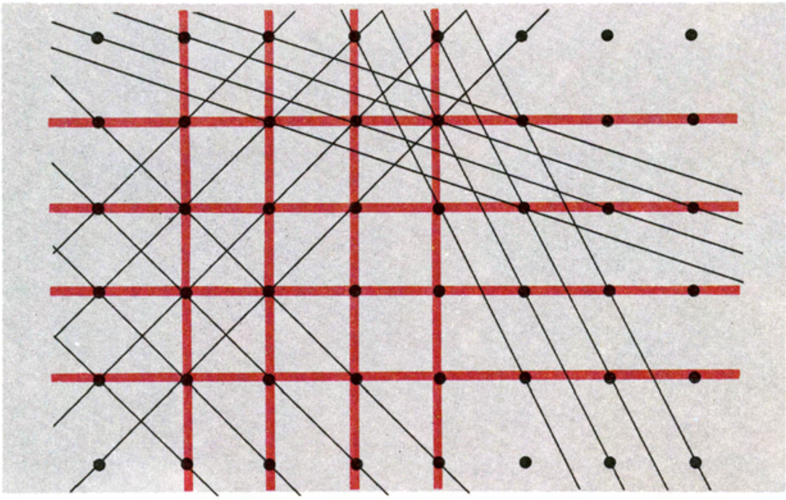


Fig. 70

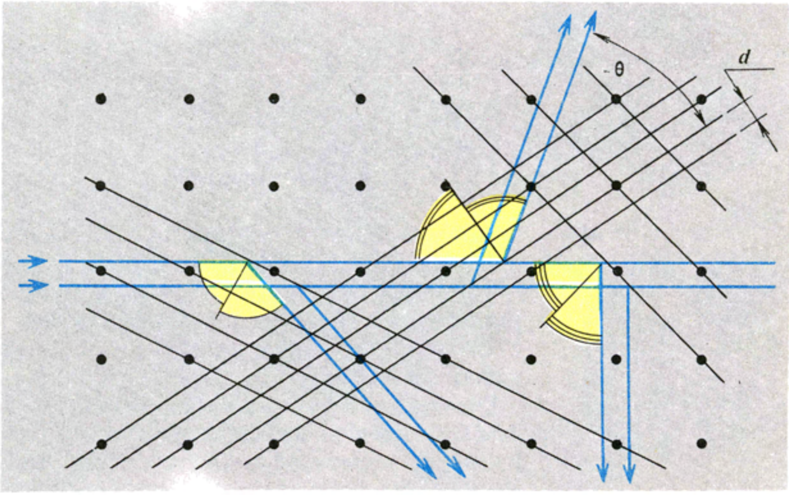


Fig. 71

that pass through the sites of the lattice. True, through lattice sites one can draw many different sets of parallel planes (Fig. 70). These sets differ in their orientation in space, interplanar spacing and density of packing of sites in the plane. Of especial interest are the most dense planes (in the figure they are shown in red). It is along these planes that a single crystal specimen normally fractures, and it is to these planes that the faces of a grown single crystal correspond. In general the faces of a unit cell are not parallel to those planes. One should not therefore expect that the shape of a single crystal specimen will coincide with the shape of a unit cell (compare Figs. 68 and 48 showing diamond).

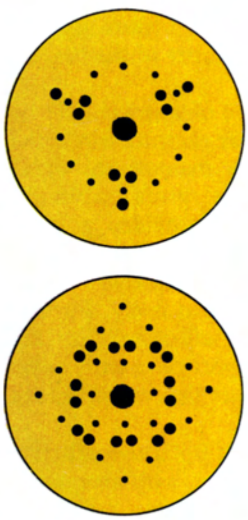


Fig. 72

The Experimental Study of Crystal Structures

Crystal structure cannot be seen even through the most powerful microscope now available. The atomic structure of a crystal is identified using the *diffraction of X-rays.* For the latter a crystal is a *diffraction grating* produced by nature.

Let us consider the simplest X-ray technique. A single crystal sample is oriented in a particular way relative to an X-ray beam. On reflecting from various sets of the parallel planes passing through the sites of the lattice, X-rays produce (on a photographic film) a picture characteristic for the given orientation of the single crystal, the so-called *Laue pattern* of the single crystal (from the name of the German physicist Laue). Each spot in the pattern corresponds to one of the reflected X-ray beams. Reflections will only be observed in directions that meet the known *diffraction condition*:

$$2d \sin \theta = n\lambda,$$

where d is the separation between neighbouring parallel reflecting planes, θ is the angle between the direction of the reflected X-ray beam and the reflecting plane (equal to a half of the angle between the directions of the reflected and initial beams), λ is the wavelength of the X-rays, and $n = 1, 2, \ldots$.

Figure 71 is a schematic representation of the reflection of X-rays from three sets of parallel planes passing through sites, supposing of course that the above condition is met.

Figure 72 is an example of the Laue pattern of *zinc blende* (ZnS) single crystal for two orientations of the sample relative to the initial beam. One of the pictures shows a four-fold rotational symmetry and the other a three-fold one. The arrangement of spots on the pictures is a tale-tell indication of the elements of symmetry of the lattice studied.

In addition to the diffraction of X-rays, some crystal studies are made using the diffraction of *electrons* and very slow *neutrons.*

The Mysteries of Water

It is common knowledge that heating reduces the density of liquids and subjecting them to higher pressure makes them more viscous. *Water*

behaves differently, however. Heating from 0 to 4 °C increases the density of water, and pressurizing it reduces its viscosity.

The mystery was unravelled after the atomic structure of water was investigated. It turned out that water molecules interact with one another in a *directed way* (just like the carbon and hydrogen atoms in the methane molecule). Each water molecule may thus form bonds with four neighbouring molecules so that their centres will form a *tetrahedron*. This is shown schematically in Fig. 73 where the balls stand for water molecules.

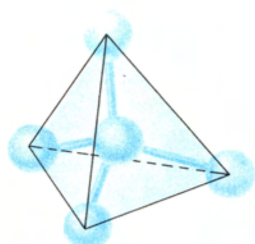


Fig. 73

This arrangement corresponds to the fairly loose, *skeleton-like* molecular structure, where each molecule has only four nearest neighbours. By way of comparison, for the dense packing of balls the number of the nearest neighbours is twelve.

Note that, unlike crystals, the molecular structure of water should be viewed as a manifestation of *short-range* order. Near each molecule the neighbouring molecules are arranged in an orderly manner, the order gradually diminishing with the distance from a given molecule.

The skeleton-like molecular structure of water provides a good explanation for its physical properties. Water increases in density when heated from 0 to 4 °C because the heating *disturbs the molecular bonds* causing the arrangement of molecules to pack more densely. Further heating reduces the density because a conflicting effect sets in: the thermal expansion of spacing between oxygen and hydrogen atoms in the water molecule. This accounts for the well-known fact that water has the *highest* density at 4 °C.

The skeleton-like molecular structure of water (near 0 °C) also explains another property of water – the drop in its viscosity with increasing external pressure. Pressurizing, just like heating, disrupts molecular bonds and thus reduces viscosity.

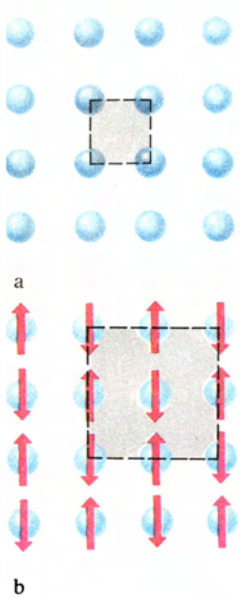


Fig. 74

Magnetic Structures

The orbiting of electrons in the field of the atomic nucleus may produce an atomic magnetic field. *Magnetic materials* may have their atomic magnetic fields ordered. So, in a *ferromagnetic* magnetized to saturation, the magnetic fields of all atoms are oriented in one direction (that of the magnetizing field), whereby the magnetic properties of the substance become especially apparent.

Also of interest is the magnetic order in a special type of magnetic material – the so-called *antiferromagnetics*, which are currently widely used in logical elements and memories of modern computers. The direction of the atomic magnetic field in those materials *alternates in a regular fashion* from one atom to the next, with the result that apart from a crystal lattice a *magnetic lattice* is present as well. For simplicity, Fig. 74*a* shows a plane square lattice, the dash lines delineating a unit cell. Figure 74*b* shows the same lattice for the antiferromagnetic. The directions of the atomic magnetic field are shown by arrows at sites, the dash lines delineating a magnetic unit cell. It is easily seen that the linear size of the unit cell is twice the size of the crystal cell.

For a real (three-dimensional) example of an antiferromagnetic, let us take *manganese oxide* (MnO), its crystal lattice is given in Fig. 75. It consists of two identical interlocked f. c. c. lattices, one of which contains manganese ions Mn^{2+} and the other oxygen ions O^{2-}. Oxygen ions have no magnetic field. The magnetic fields of manganese ions in the planes are represented in the figure by the same colour (for example, red), they are oriented in the same way, whereas the magnetic fields of manganese ions belonging to the planes of different colour are oriented in

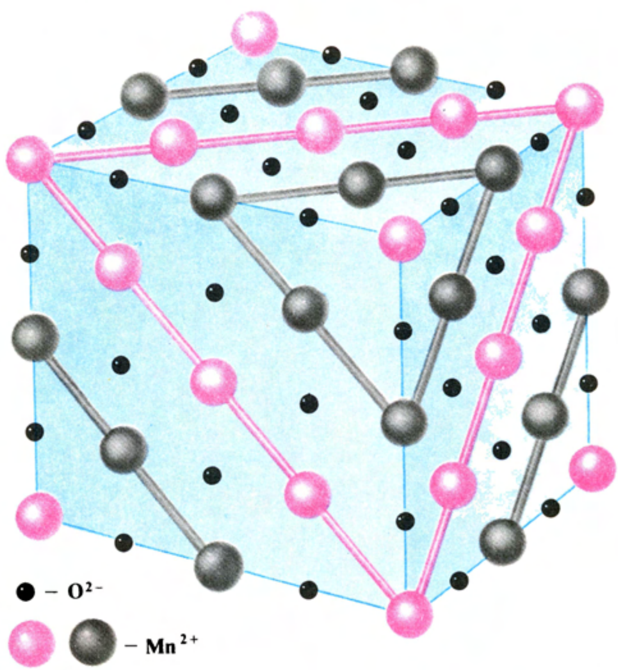


Fig. 75

opposite directions (the directions of magnetic fields are perpendicular to the planes chosen).

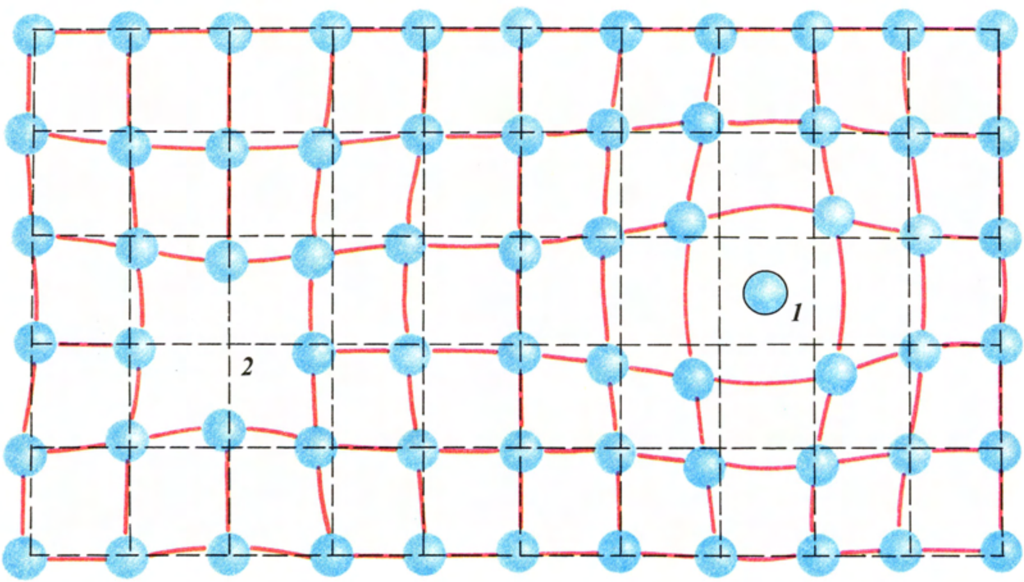
Order and Disorder

In the surrounding world *order and disorder are inseparable.* Whatever the degree of order of a given atomic structure, it still has some elements of disorder. Specifically, this is true of the atomic structure of crystals.

It is to be noted first of all that atoms are by no means fixed at lattice sites, but are instead involved in *thermal vibrations* near those sites, the amplitude being the larger the higher the temperature. Thermal vibrations make individual atoms leave their sites and wander (diffuse)

about the crystal. In the lattice some unoccupied sites (so-called *vacancies*) appear. Both vacancies and atoms at gaps between sites (interstices) will obviously distort the geometry of the lattice by influencing the arrangement of neighbouring atoms (Fig. 76). What is more, any real crystal may have some foreign atoms, the so-called *impurities*. These atoms may be at interstices, but may also substitute for lattice atoms by "driving them away" from their places.

Substantial disorder in the lattice is caused by so-called



1 — interstitial atom

2 — vacancy

Fig. 76

dislocations – violations of the regular arrangement of atomic planes. Some idea of dislocations may be obtained in Fig. 77, which represents the so-called *edge* dislocation.

The presence of defects in a crystal lattice is the decisive factor influencing the *strength* and *plasticity* of a material. At present, special techniques enable filament crystals to be grown, such that their lattice is essentially defect-free. The strength of such crystals may be as high as 10^{10} Pa, hundreds of times larger than for conventional crystals. According to modern thinking, the plasticity of a material is controlled by the *migration of defects*, above all dislocations, over the sample. Interestingly, if the density of the defects grows, they eventually begin to hinder the migration with the result that the plasticity of a material drops. This is what occurs in certain kinds of processing (forging, annealing, etc.).

Consequently, a material can be strengthened by two *opposite* ways – either by preventing defect formation, or by hindering the migration of defects about the sample (that is, by increasing the density of defects). The first technique means growing defect-free crystals, and the second one a special processing of materials.

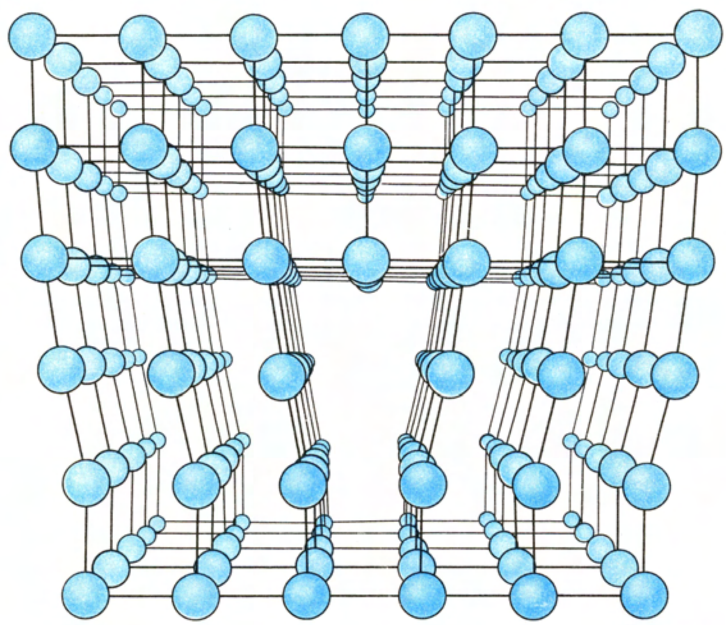


Fig. 77

7 Spirality in Nature

Pasteur was more right than many of his colleagues suspected when he wrote eloquently of left-right asymmetry as a key to the mystery of life. At the heart of all living cells on earth are right-handed coils of nucleic acid. This asymmetric structure is surely the master key of life.

M. Gardner

Mirror asymmetry (also called *left-right asymmetry*) widely occurs in nature and is of principal importance for living things. A characteristic example of a mirror-asymmetric object is a *helix* or a *spiral*. This explains why left-right asymmetry is often referred to as spirality. Also used is the term handedness. Recall that the *hand* is an example of a mirror-asymmetric object.

The Symmetry and Asymmetry of the Helix

The figure in Fig. 78 appears to be unmoved if turned about axis OO through 60° and translated along the same axis through distance a. The axis is a six-fold *helical axis* with translation period a. *Helical symmetry* is a symmetry relative to a combination of a turn and a translation along the rotation axis. An example of an object featuring helical symmetry is the spiral staircase.

For an object having helical symmetry we can draw a *helical line*, or spiral (the dash line in Fig. 78). The helical line can be constructed as follows. Cut a right-angled triangle out of paper (ABC in Fig. 79). Take a circular cylinder and glue to its surface the leg BC of triangle ABC so that the leg coincides with the generatrix of the cylinder surface. Next wrap the triangle around the cylinder tightly pressing the paper to the cylinder surface, the hypotenuse AB will give the helical line. Two methods of turning the triangle around the cylinder are possible, both of which are shown in Fig. 79. One of them produces a *left-handed* and the other a *right-handed* spiral.

The type of a helical line is identified rather simply. Let us mentally move along a helical line, the motion will have two components – *along* the helical axis and *around* the axis (the straight and circular arrows in Fig. 79). Let us place the observer so that the structure moved along the helical axis *away from him*. If in the process the circular motion is *clockwise*, then the helical line is termed *right-handed*, and if *counterclockwise*, it is called *left-handed*. In other words, if a point moving away from the observer in a helical line appears to him to turn clockwise, the spiral is a right-handed one, if counterclockwise, a left-handed one. When reflected in a mirror, a left-handed spiral will become a right-handed one, and vice versa. The left- and right-handed spirals form a pair of *enantiomorphs*.

Note that speaking about the helical line or the screw, the term "spiral" is often used. It will be remembered that we have here a *spatial* spiral.

Helices in Nature

We use helices widely in technology. It is perhaps of interest that helices are also encountered in nature. Some examples of natural helices are shown in Fig. 80: (a) the tusk of the narwhal, a small cetacean animal endemic of northern seas, is a left-handed helix, (b) the shell of a snail*, (c) the umbilical cord of a new-born is a triple left-handed spiral formed by two veins and one artery, (d) horns of the Pamir sheep are enantiomorphs (one horn is left-handed and the other right-handed spiral).

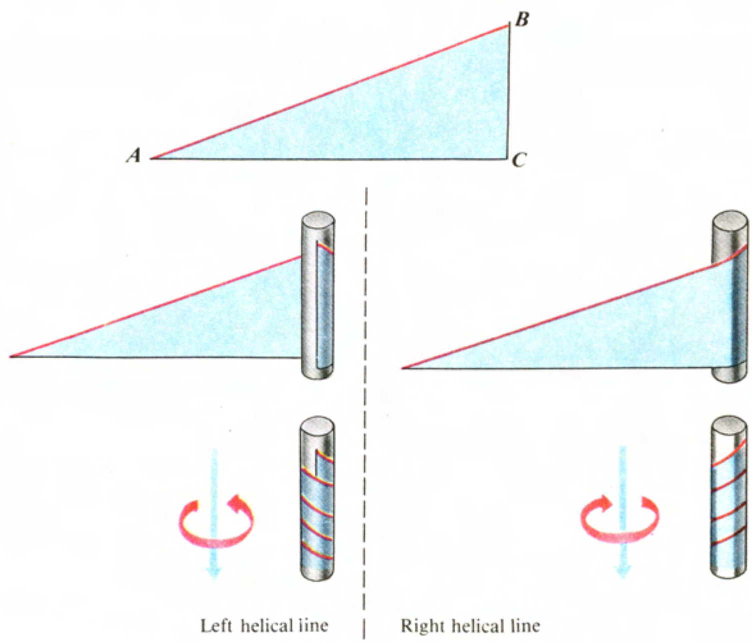


Fig. 78

In plants we may find numerous signs of helical symmetry in the arrangement of leaves on the stalk, branches of the stem, the structure of cones, some flowers, and so on. Creeping plants are remarkable helices. Examples are known of entangled creeping plants with different handedness, for example *bindweed* and *honeysuckle*. This gives rise to bizarre arrangements, which repeatedly attracted poets. Crystal lattices, as a rule, have mirror symmetry. But there exist also mirror-asymmetric lattices, some of them are characterized by a helical structure. An example of a twisted crystal lattice is *quartz*. Its base is a *tetrahedron* with

* Each type of shell has a certain spirality. Any "freaks", which are encountered from time to time, having the opposite spirality, are especially valued by collectors.

a silicon atom at the centre and oxygen atoms at its vertices. Along the main crystal axis the tetrahedra lie along a *helical line*. The quartz lattice may be twisted either to the left or to the right. Therefore, there exist two enantiomorphic forms of quartz. The left and right single crystals of quartz are represented in Fig. 81. It is easily seen that one is a mirror image of the other.

Another realm of natural screws is the world of "living molecules", those molecules that play an important role in biological processes.

Fig. 79

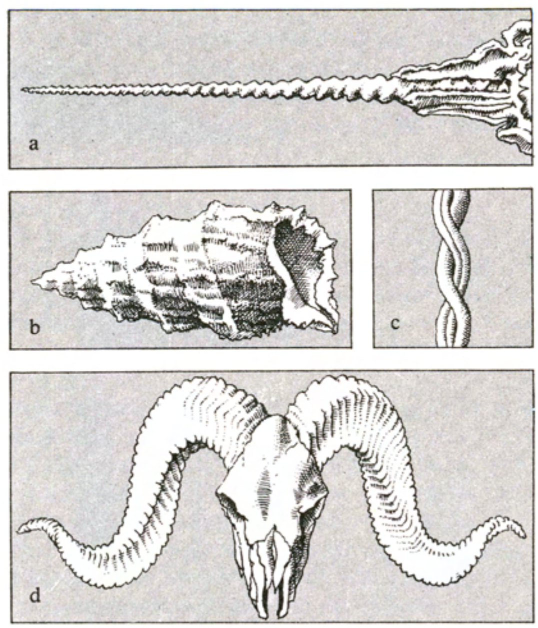


Fig. 80

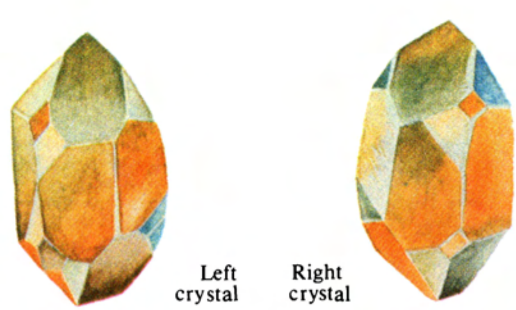


Fig. 81

These molecules include, above all, *protein molecules*, the most complex and numerous of all the carbon compounds. In man there are up to 10^5 types of proteins. All the human organs, including bones, blood, muscles, sinews, and hair, contain proteins. The many ferments and hormones are proteins as well.

The *protein molecule* contains atoms of carbon, hydrogen, oxygen, and nitrogen. Each molecule has an enormous number of atoms, of the order of 10^3-10^6. Each giant molecule contains many links (amino acids) connected into a chain. The frame of the chains is twisted as a right-handed spiral. In chemistry it is called *Pauling's alpha-spiral* (after the prominent American chemist Linus Pauling). The molecules of sinew fibres are triple alpha-spirals. Alpha-spirals twisted repeatedly with each other form molecular helices such as those found in hair fibres, in horny substance, and so on.

Especially important in living nature are the molecules of *deoxyribonucleic acid* (DNA), which are bearers of hereditary information in living organisms. The DNA molecule has the structure of a *double right-handed spiral* which, in a sense, is the main natural helix. Let us take a closer look at the structure of that molecule.

The DNA Molecule

A schematic diagram of the DNA molecule is given in Fig. 82. The molecule consists of many links called *nucleotides*, the links being connected into two chains (in the figure the nucleotides are shown by red rectangles). Each nucleotide contains a *sugar* molecule, a phosphoric acid molecule (*phosphate*), and a molecule of a nitrogen-containing compound (*nitrous compound*). The nitrogenous bases of two nucleotide strands are linked by hydrogen bonds shown by dash lines in Fig. 82. The structure in the figure looks like a ladder, in which the vertical elements are *sugar-phosphate chains*, and the rungs are *pairs of nitrogenous bases*.

There are four nitrogenous bases: *adenine* (A) and *guanine* (G) (purines), *thymine* (T) and *cytosine* (C) (pyrimidines). Each rung contains either adenine and thymine (A—T or T—A), or guanine and cytosine (G—C or C—G). Figure 83 represents the structural formulae of the adenine-thymine pair and the guanine-cytosine pair, which enter the DNA molecule. There are no combinations of adenine with guanine or thymine with cytosine.

Considering the above, the "ladder" schematically representing the structure of DNA assumes the form shown in Fig. 84. The arrangement of the pairs AT, TA, GC, and CG along the "ladder" is the *genetic code* of a living organism. Despite the fact that there exist only four types of rungs, the enormous number of these rungs on the ladder enables DNA to include all hereditary information.

This information is retained when cells are reproduced. The DNA molecule divides into two halves (along the red line in Fig. 84), each of which is essentially a sugar-phosphate chain with nitrogenous bases arranged normally to the chain. Since each base may only bond to a *specific* base (A to T, T to A, G to C, C to G), then each of the halves will

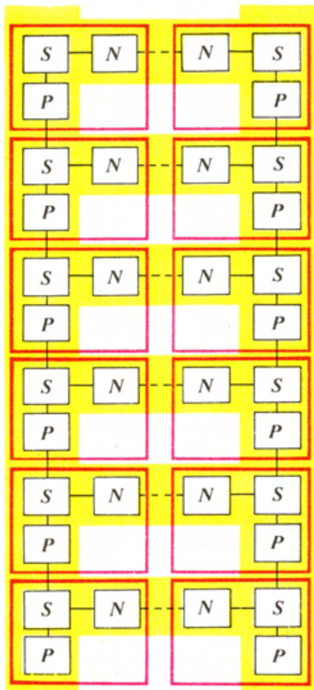


Fig. 82

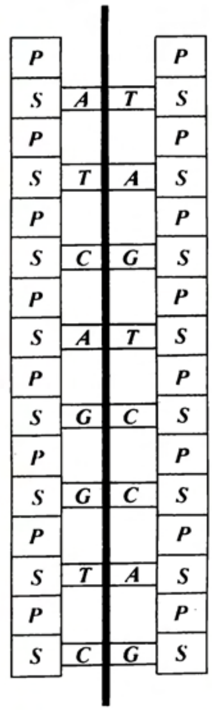


Fig. 84

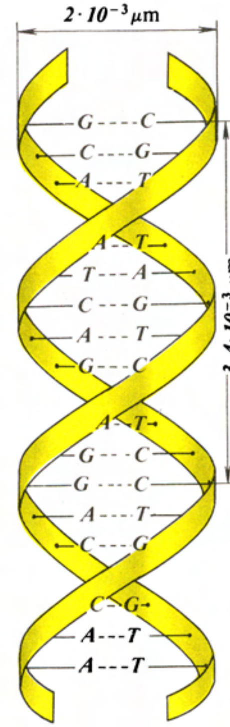


Fig. 85

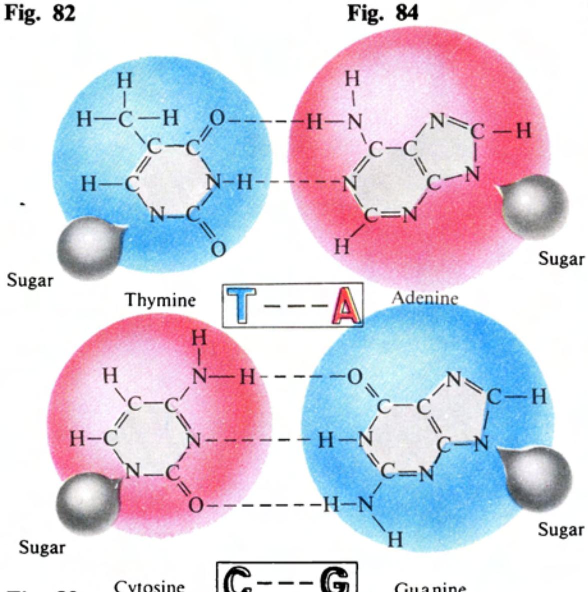


Fig. 83

build up to form a molecule that is a complete replica of the initial DNA.

In passing from the schematic representation of the DNA molecule to its real *spatial* structure, it is necessary to take into account the fact that each sugar-phosphate chain represents a right-handed spiral, so that on the whole the DNA molecule appears as a *double right-handed spiral* and looks like a spiral staircase, rather than a ladder (Fig. 85).

It is worth noting that all the spirals in the DNA molecules in man are right-handed. Among the prodigious wealth of them not a single left-handed one is to be found. Note that the DNA strands in one cell may be up to a metre in length. All in all, there are 10^{11} kilometres of DNA in man.

The structure of DNA was discovered in 1953 by a team of scientists that included the American Watson and the Englishmen Crick and Wilkins. This discovery is justifiably regarded as one of the most remarkable contributions to biology in the 20th century.

The Rotation of the Plane of Light Polarization

Some media possess a fascinating property: when a plane-polarized light beam is passed through them, the light polarization plane turns by a certain angle. Such media are termed *optically active*. The media may be *levorotary* or *dextrorotary*. Suppose that the beam strikes our eye. If the polarization planes turn *clockwise*, the medium is called *dextrorotary* (Fig. 86*a*); if *counterclockwise*, *levorotary* (Fig. 86*b*).*

We are not going to consider the nature of the phenomenon of polarization plane rotation. We will only note that an optically active medium must possess a left-right asymmetry, which dictates the rotation of the polarization plane in one direction or the other.

An example of optically active media is the quartz crystal. The direction of the turn of the polarization plane depends on to which enantiomorphic variety a given crystal belongs. Dextrorotary crystals are generally called right-handed, and levorotary crystals are called left-handed ones. The optical activity of quartz is associated with the left-right asymmetry of the lattice. If a quartz crystal is dissolved in a liquid, no rotation of the plane of light polarization is observed.

It would seem that the presence of a left-right asymmetry of crystal structure is a necessary condition for a polarization plane to turn. You can imagine the surprise of the prominent 19th century physicist Jean-Baptiste Biot when he discovered optical activity in aqueous solutions of some organic compounds, for instance, *sugar* and *tartaric acid*. It thus followed that the left-right asymmetry could be associated not only with the structure of a medium as a whole but also with the structure of molecules in the medium. This gave rise to the terms "left" (levorotary) and "right" (dextrorotary) molecules.

---

* Note that the "left-handed" ("right-handed") combination of the straight and circular arrows in Fig. 86 differs from the appropriate combination in Fig. 79. This implies that in a *levorotary* medium the polarization plane turns following in fact a *right-handed* helix, and in a *dextrorotary* medium it turns following a *left-handed* helix.

Left and Right Molecules. Stereoisomerism

In Chapter 6 we discussed the molecules of methane and methyl alcohol (see Fig. 59). Both molecules are identical to their mirror images. This is only natural since they feature mirror symmetry (six planes of symmetry in the methane molecule and three planes of symmetry in the methyl alcohol molecule).

Let us now substitute radical CH_3 for one of the three identical hydro-

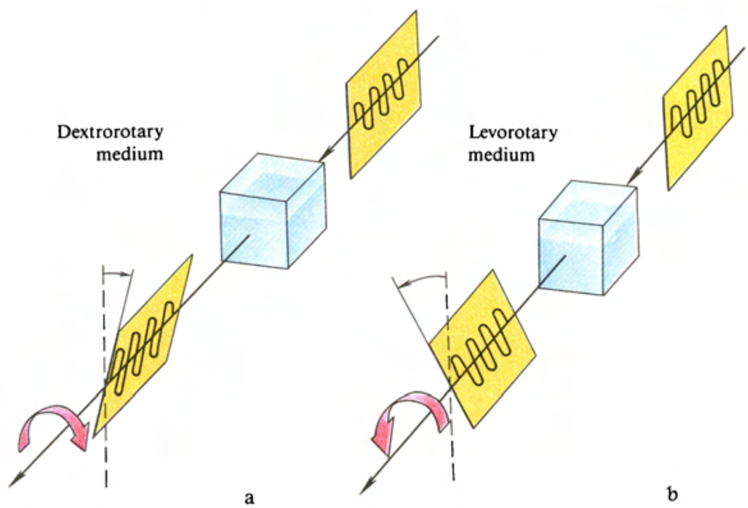
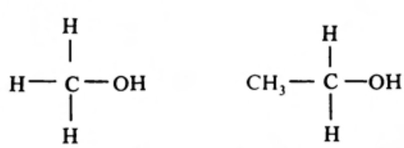


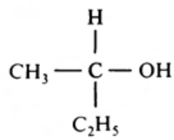
Fig. 86

gen atoms in the molecule of methyl alcohol, that is, pass from the structural formula



This is the formula of *ethyl alcohol*. The tetrahedral model of this molecule is provided in Fig. 87*a*, and the model of its mirror image in Fig. 87*b*. It appears that in this case too the mirror image is identical to its original molecule: to obtain a mirror image we will have to turn the molecule of ethyl alcohol by 180° about axis *AB* (see the figure). In other words, the molecule of ethyl alcohol (just like those of methane and methyl alcohol) has no enantiomorphic versions. This is because despite the gradual reduction of the symmetry of molecular tetrahedron (in passing from methane through methyl alcohol to ethyl alcohol), it still remains mirror-symmetrical. It is easily seen that the molecule of ethyl alcohol has a plane of symmetry (plane *ABD* in Fig. 87*a*).

A different situation emerges if we take, for example, the molecule of *butyl alcohol*



The spatial tetrahedral model of the molecule is presented in Fig. 88*a*, and its mirror image in Fig. 88*b*. The molecule has no plane of symmetry, it is *mirror-asymmetrical*. Therefore, both it and its mirror image are enantiomorphs, and no turns can effect a coincidence of the molecules shown in Fig. 88. One of the molecules can be termed "left", and the other "right".

Thus, if the spatial structure of a molecule excludes planes of symmetry, then it can have two forms, which are enantiomorphs. These forms are called *stereoisomers*.

Stereoisomerism is a manifestation of left-right asymmetry in the world of molecules. Stereoisomers are molecules that in addition to the same chemical composition have the same geometrical shape, the same structural elements, and the same inner bonds. At the same time they are *different* molecules. As different as, say, left and right shoes. The existence in nature of left- and right-handed molecules was suggested by observations of the rotation of polarization planes.

A special case among stereoisomers are, obviously, molecules with spiral structures but differing in handedness.

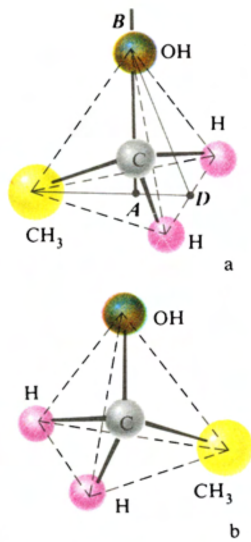


Fig. 87

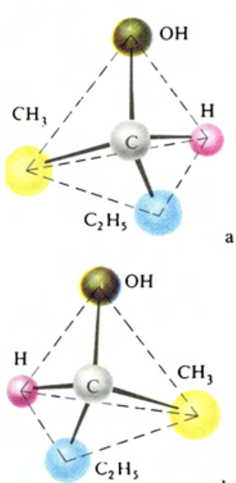


Fig. 88

The Left-Right Asymmetry of Molecules and Life

Studies of the optical activity of solutions of organic compounds, initiated by Biot, were carried on by the famous French scientist Louis Pasteur. He came to the conclusion that if in *inorganic* nature left- and right-handed molecules occur with equal frequency, in *living* organisms mirror-asymmetrical molecules occur only in one enantiomorphic form. Pasteur speculated that it is here that the *boundary* between organic and inorganic nature lies. Further scientific evidence supported Pasteur's conjectures.

Recall that the alpha-spiral, which determines the structure of protein molecules, is invariably a right-handed spiral. The double spiral of the DNA molecule is also right-handed. Various mirror-asymmetrical molecules that enter the compositions of cells are, as a rule, represented either only by left-handed or only right-handed stereoisomers. So, amino acid molecules in proteins are always left-handed.

This all goes to prove that on the molecular level, a *living organism is characterized by distinct left-right asymmetry*. An organism is constructed out of "helices", so to speak, some being only right-handed and others only left-handed.

One manifestation of that highly interesting thing is that left- and right-handed stereoisomers of a substance exert different actions of the

human body. Man consumes in food those stereoisomers that *correspond to the nature of his own asymmetry*. In a number of cases modern chemistry is able to obtain mirror-reflected stereoisomers. And then we can observe unexpected reactions of the human organism to them. So, the "reflected" stereoisomer of vitamin C exerts essentially no effect on the body. The "reflected" form of nicotine (never to be found in tobacco) is much less offensive. Right-handed aspartic acid is sweet, left-handed is tasteless. Minor additions of right-handed phenylalanine to food have no unpleasant implications, whereas additions of the left-handed form produce drastic metabolic disturbances (phenylketonuria) accompanied by mental disorders.

Suppose now that after a long space journey a man stepped onto an unknown planet. Suppose further that the planet is very much like the Earth (in its atmospheric composition, climate, landscape, plants, etc.). And so, the astronaut holds in his hand a fragrant apple just picked from an extraterrestrial appletree. But should he eat the apple? The enantiomorphic form of organic compounds in it is unknown. It is quite possible that an innocent-looking apple may appear to be biologically poisonous for an Earthling. In other words, mirror-asymmetrical molecules of a foreign plant world may turn out to be *incompatible* with the mirror-asymmetrical human organism, just as a *left-threaded* nut is incompatible with a *right-threaded* bolt.

The beautiful children's book *Through the Looking Glass* by Lewis Carroll contains a scene that today has deep scientific meaning. About to pass through the looking glass into the looking-glass house, Alice asks her kitten: "How would you like to live in Looking-Glass House, Kitty? I wonder if they'd give you milk in there? Perhaps looking-glass milk isn't good to drink?" As a matter of fact, milk includes many mirror-asymmetrical compounds, such as fats, lactose (milk sugar), and proteins. On passing over from the conventional world to the looking-glass world, all the asymmetrical molecules should have turned from some stereoisomers to other ones. Therefore, it is highly unlikely that the looking-glass milk would have been wholesome for the kitten. Although, if we were consistently to follow the situation described by Carroll in his book, we are to assume that in the looking-glass world both Alice and the kitten themselves would turn into their respective mirror images. And in that case the looking-glass milk would, of course, be for them as palatable and useful as the conventional, "unreflected" milk had been before.

We will conclude the chapter with the words of M. Gardner from his book *The Ambidextrous World*: "One of the most remarkable and least mentioned characteristics of life as we know it is the ability of an organism to take compounds from its immediate environment, many of which are symmetrical in their molecular structure, and to manufacture asymmetrical carbon compounds that are right- or left-handed. Plants take symmetrical inorganic compounds such as water and carbon dioxide and from them manufacture asymmetrical starches and sugars. The bodies of all living things are with asymmetrical carbon molecules, as well as the asymmetrical helices of proteins and nucleic acids."

Part TWO

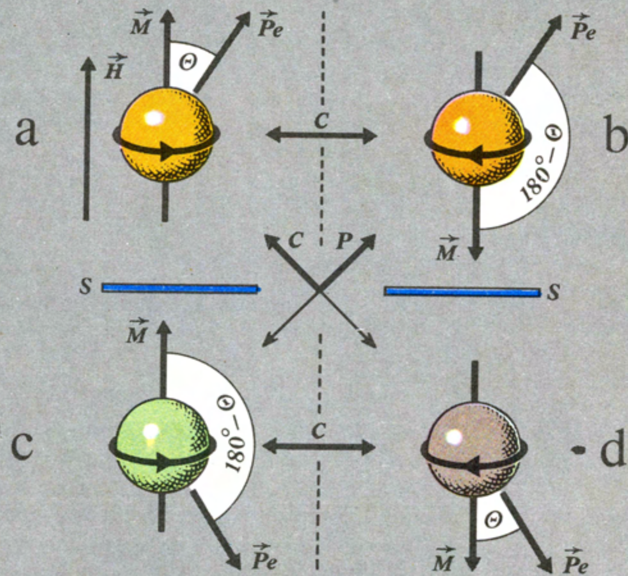
Symmetry at the Heart of Everything

There is a steady system in all,
Full consonance in nature;
It is only in our phantom freedom
That we sense discord with her.

F. Tyutchev
(Translated by Jesse Zeldin)

To see a World in a Grain of Sand
And a Heaven in a Wild Flower
Hold Infinity in the palm of your hand
And Eternity in an hour.

W. Blake



a
b
c
d
$\vec{M}$
$\vec{H}$
$\vec{Pe}$
Θ
$180° - \Theta$
C
P
S

8 Symmetry and The Relativity of Motion

The laws governing natural events are independent of the state of motion of the reference frame in which these events occur if the frame travels without acceleration.

A. Einstein

The concept of symmetry is not confined to the symmetry of objects. It also covers *physical events* and *laws* that govern them. The symmetry of physical laws resides in their *unchangeability*, or rather *invariance*, under one or another of transformations related, for example, to the conditions under which the phenomenon is observed.

Scientists got interested in the issue of symmetry in the laws of physics in connection with the studies that had led to the development of special theory of relativity. The symmetry here is the symmetry (invariance) of the laws of physics *under transition from one inertial reference frame to another inertial reference frame*, or the symmetry in relation to uniform motion.

The Relativity Principle

Suppose that a railway carriage moves smoothly and uniformly. You travel in the carriage and want to find out whether it is moving or it is at rest. Could this be done without looking out of the window? The reader must know the answer. In that case it is in principle impossible to find this out because all the physical processes occur in the *same manner* in a carriage at rest and in a carriage in uniform motion.

This is known as the *principle of relativity* for inertial reference frames. Recall that a reference frame is said to be inertial if a body in it moves uniformly when not exposed to the action of external forces. Any two inertial reference frames are in uniform motion relative to each other. In the above example of the train carriage one inertial frame is related to bodies at rest on the ground, the other is related to the carriage, which moves uniformly and rectilinearly. We can here ignore the rotation of the Earth and its revolution about the Sun.

The principle of relativity can be formulated as follows. *Any process in nature occurs in the same manner in any inertial reference frame; in all inertial frames a law has the same form.*

As applied to mechanical phenomena the relativity principle was established by Galileo. The principle was generalized to apply to all processes in nature, including *electromagnetic* ones, by the foremost physicist of the 20th century Albert Einstein (1879–1955), the father of relativity theory. Drawing on the very essence of the relativity principle, Einstein postulated that the *velocity of light in a vacuum must be the same in all inertial reference frames.* Einstein's discovery was a great breakthrough in science, since it subjected all the age-old views of space and time to an overhaul.

The Relativity of Simultaneous Events

It follows from the invariance of the speed of light with respect to translations from one inertial reference frame to another that two spatially separated events that are *simultaneous* in one frame may be *nonsimultaneous* in the other. Let us take a simple example.

Consider two inertial frames of reference xyz and $x'y'z'$. Let frame $x'y'z'$ be travelling relative to frame xyz along x- and x'-axes with a speed v (Fig. 89). In frame $x'y'z'$ there are a light source A and two light

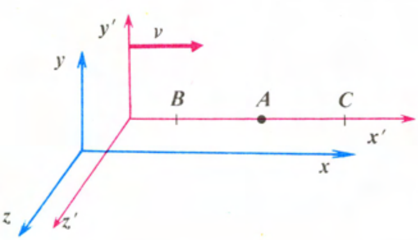


Fig. 89

detectors, B and C, that lie at equal distances from A along the x'-axis. Light source A sends out two light impulses simultaneously, one in the direction of B, the other in the direction of C. Since $AB = AC$ and both signals travel with the same speed, the observer in $x'y'z'$ will see detectors B and C operate *simultaneously*.

Let's turn now to the observer in xyz. In this frame of reference the light signal that travels to the left will have to cover a smaller distance from production to registration than the signal that travels to the right. The speed of light in xyz and $x'y'z'$ being the same, for the observer in frame xyz detector B will operate *earlier* than C.

The Lorentz Transformations

We will stay with our inertial frames of reference xyz and $x'y'z'$ shown in Fig. 89. Let an event occur at a time t at a point x, y, z in the frame xyz. In the frame $x'y'z'$ the same event occurs at the time t' at the point x', y', z'. The space and time coordinates of the event in the frames xyz and $x'y'z'$ are related by

$$\left.\begin{aligned} x' &= \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}; \\ y' &= y; \\ z' &= z; \\ t' &= \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}, \end{aligned}\right\} \qquad (1)$$

where c is the speed of light in empty space. These relationships are called the *Lorentz transformations* (from the Dutch physicist Hendrik Lorentz who first derived them).

The symmetry of physical laws with respect to changes from one inertial reference frame to another are mathematically expressed in that the relevant mathematical expressions must conserve their form if x, y, z, t in them are replaced by x', y', z', t' according to (1). Put another way, the mathematical expressions of physical laws must possess *symmetry with respect to the Lorentz transformations.* We will illustrate this important statement with reference to the physical law that states the constancy of the speed of light in all the inertial coordinate systems.

Suppose that a light signal originates from point $x = y = z = 0$ at $t = 0$ along the x-axis in the system xyz. At a time t it will be registered at the point $x = ct$, $y = z = 0$. If the speed of light is the same in all the inertial systems, then, substituting $x/t = c$ into (1), we arrive at $x'/t' = c$. We will now see that this is so. Dividing the first of (1) by the last one gives

$$\frac{x'}{t'} = \frac{x - vt}{t - \frac{vx}{c^2}} = \frac{\frac{x}{t} - v}{1 - \frac{vx}{c^2 t}} = \frac{c - v}{1 - \frac{v}{c}} = c.$$

Using (1), we can easily demonstrate the relative nature of the simultaneousness of events. Let two events have in frame xyz the space-time coordinates x_1, t_1 and x_2, t_2 (in this case we can neglect the coordinates y and z), and in frame $x'y'z'$ the coordinates x_1', t_1' and x_2', t_2', respectively. From (1), we get

$$t_2' - t_1' = \frac{(t_2 - t_1) - \frac{v}{c^2}(x_2 - x_1)}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Suppose that the events are simultaneous in frame xyz, but at different spatial points. This means that $t_2 = t_1$, $x_2 \neq x_1$. It is easily seen that in this case $t_2' \neq t_1'$; in other words, in frame $x'y'z'$ the events are not simultaneous.

If the speed of the relative motion of xyz and $x'y'z'$ (speed v) is *much lower* than the speed of light, then relationships (1) simplify to yield

$$\left.\begin{aligned} x' &= x - vt; \\ y' &= y; \\ z' &= z; \\ t' &= t. \end{aligned}\right\}$$

These relationships are known as the *Galilean transformations*. They reflect the principle of relativity in classical mechanics due to Galileo. From the viewpoint of the special theory of relativity, the Galilean transformation equations are a *special case* of the Lorentz transformation equations, which is valid at $v \ll c$.

It is worth noting that, while recognizing the relative nature of spatial coordinates ($x' = x - vt$), the Galilean transformation equations at the same time assumed that *time is absolute* ($t' = t$). The notion of absolute time is deeply rooted in the human mind. We are in the habit of thinking that the phrase "the event is occurring *now*, at this moment" has the same sense for all reference systems and for the whole of the Universe.

Physically, the fundamental difference between the Galilean and Lorentz transformations lies in the fact that in the first case we ignore the finiteness of the propagation of light velocity signals, whereas in the second case we take it into account. When dealing with relatively slow motions the approximation is quite justified. As follows from (1), with finite speeds we have to *forego the absolute nature of time and consider the spatial and temporal coordinates jointly.* This is a consequence of the *symmetry of physical laws with respect to changing from one inertial reference frame to another, which manifests itself when we take into account the finiteness of the speed of light.*

The Relativity of Time Periods

Suppose two events occur at the same point in frame $x'y'z'$ separated by a time τ'; in symbols: $x_2' - x_1' = 0$, $t_2' - t_1' = \tau'$. Suppose further that frame $x'y'z'$ is connected, say, with a spacecraft travelling at a velocity v relative to the Earth, and the above-mentioned *events* are "the astronaut left his chair" and "the astronaut returned to his chair". The frame $x'y'z'$ is called the *rest frame* for these events, since they are, as it were, at rest, i. e., occur at the same point in space. The time period between the two events in the rest frame is called the *proper time.*

Let us now turn to a frame xyz connected with the Earth. The above events, if considered in frame xyz, i.e. from the Earth, will occur at *different* spatial points: x_2 and x_1. The events are separated by the time span $\tau = t_2 - t_1$ by the clock of the terrestrial observer. Using (1), we can easily find that

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}; \qquad t = \frac{t' + \frac{vx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Hence

$$\tau = t_2 - t_1 = \frac{(t_2' - t_1') + \frac{v}{c^2}(x_2' - x_1')}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{\tau'}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Consequently, the time between two events depends on the choice of the reference frame. This time is minimal in the rest frame for the given events

(proper time). The time increases by a factor of $(1 - v^2/c^2)^{-1/2}$ in the frame that moves at a speed v relative to the rest frame.*

Suppose that the speed of the ship is fairly close to the speed of light; for example, $v/c = 0.9999$ (maybe such ships will be available in the future). In this case $\tau \approx 70\tau'$. The astronaut has only left his chair for 20 minutes, and on Earth the clock showed all of 24 hours.

However, it can easily be surmised that this situation is *reversible*. You, on Earth, will take half an hour to read this section, and on the spaceship this time span will again be 24 hours. The situation will only be irreversible if the ship returns back to Earth in the long run. But this is a separate topic that lies beyond the scope of the book.

The Speed in Various Frames

Before leaving the subject of the special theory of relativity we will consider how the *speed of a body* changes when we go over from one inertial frame to another. We can easily feel the contradiction between the classical composition of speeds and the postulate that speed of light is constant in all inertial frames. The special theory of relativity requires that the classical rule be replaced by another, more general one.

Let a body move in a frame xyz uniformly with a speed V along the x-axis, and let the speed of the body be V' in a frame $x'y'z'$ that moves with a speed v relative to xyz (see Fig. 89). Considering that $V = x/t$ and $V' = x'/t'$, from (1) we directly obtain the rule

$$V' = \frac{V - v}{1 - \frac{Vv}{c^2}} \quad \text{or} \quad V = \frac{V' + v}{1 + \frac{V'v}{c^2}}.$$

If $v \ll c$, then the denominator becomes unity, and so we arrive at the classical rule: $V = V' + v$.

---

* In connecting the rest frame with a spaceship, we consider the motion of the Earth relative to the ship. It will be recalled that it makes no sense to find out which of the two inertial frames is *actually* moving and which is at rest, since there only exists *relative* motion.

9 The Symmetry of Physical Laws

When learning about the laws of physics you find that there is a large number of complicated and detailed laws, laws of gravitation, of electricity and magnetism, nuclear interactions, and so on. But across the variety of these detailed laws there sweep great general principles which all the laws seem to follow. Examples of these are the principles of conservation, certain qualities of symmetry....

R. Feynman

The symmetry of the laws of physics with respect to the Lorentz transformations (or relative to changing from one inertial frame to another) is one of the most striking examples of this kind of symmetry. Also there are other forms of symmetry of physical laws.

Symmetry Under Spatial Translations

On a wide board we will install several physical devices: a mathematical pendulum, a pair of communicating vessels, an electric circuit consisting of a battery, a switch, connecting wires, and three identical ammeters, two of which are connected in parallel. Let us now check that our elementary physical laboratory functions in full conformity with the laws of physics: the period of the pendulum's swing is controlled by its length according to the formula $T = 2\pi\sqrt{l/g}$, the level of water in both vessels is the same, the reading of either of the parallel-connected ammeters is one half that of the third ammeter. Let us now transfer our laboratory to another room. Clearly, it will function there precisely as before. This simple example is a graphic illustration of the *invariance of the laws of physics under spatial translations.*

Just to get an idea of how important this type of symmetry is, let us try and imagine what would happen if physical laws were changed by spatial translations. You move to another flat and to your surprise find that your TV set, which has functioned perfectly, now will not work. You shift a clock on your table and it either stops or goes wrong. Swimmers perform differently in different swimming pools since the water resistance changes from place to place. Results obtained in an experiment carried out in Moscow University cannot be checked at Oxford using exactly the same equipment. And so on and so forth. It is easily seen that the demise of the symmetry of physical laws with respect to spatial translations immediately leads us to a picture of some absurd, unreliable world.

To be sure, when speaking about translational symmetry we should be aware of the fact that relative translations of objects may affect the intensity and nature of their interaction. Naturally, transferring a pendulum to the Moon would change its period. Moving a clock may indeed render it inoperative if it is put on a strong magnet. Shifting a TV set farther away from a transmitting aerial will adversely affect its operation or put it out of the aerial's range. In shifting something from one point in space to another, one should consider the *environment*, or rather the degree of its influence on the functioning of the device in question.

Feynman writes: "It is necessary, in defining this idea, to take into account everything that might affect the situation, so that when you move the thing you move everything."

The invariance of physical laws under spatial translations can be shown referring to the example of the *law of gravitation:*

$$F = G\,\frac{m_1 m_2}{R^2}.$$

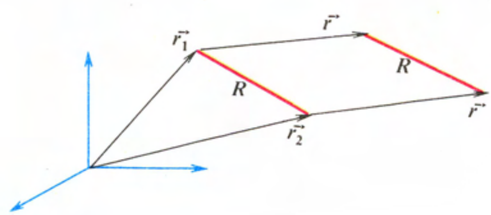


Fig. 90

To be more specific, we will assume that m_1 and m_2 are the solar and terrestrial masses respectively. Then R is the separation between the centres of the celestial bodies, and F is their mutual attraction. The constant G is known as the gravitational constant. Suppose now that the positions of the Earth and Sun relative to some point within the Galaxy chosen as the origin of coordinates are given by the position vectors $\vec{r}_1$ and $\vec{r}_2$ respectively. The attracting force F for a given time can be found from the law of gravitation, where we put $R = |\vec{r}_2 - \vec{r}_1|$ (Fig. 90). After a *spatial translation* we will have to add to $\vec{r}_1$ and $\vec{r}_2$ the translation vector $\vec{r}$ (see the figure). It is easily seen that this does not change R in the expression

$$R = |(\vec{r}_2 + \vec{r}) - (\vec{r}_1 + \vec{r})| = |\vec{r}_2 - \vec{r}_1|.$$

It follows that a translation does not affect the law of gravitation.

The invariance of the laws of physics under spatial translations is normally described by the term the *homogeneity of space.*

Rotational Symmetry

The laws of nature are invariant not only under *translations* but also under *rotations* in space. It is immaterial to a TV set in which direction its screen faces (although this is of consequence for the viewers). Experimental results are not influenced by the fact that the experimental set-up is oriented to north, not to east (save for geophysical measurements). Together with our Earth we are all involved in the complex motion around the Sun, this motion is a combination of rotations and translations. Should physical laws not be invariant under rotations and translations in space, they would change in time with a period of a year. In that case, the finding obtained in one or another of

physical experiments would, in addition, depend on the month in which the experiment would be carried out.

The invariance of the laws of nature under rotations is normally described by the term the *isotropicity of space*.

The notion of isotropicity of space has not come easily to humanity. In ancient times it was widely believed that the Earth is flat and so the vertical direction is an absolute one. When the concept of the sphericity of the Earth took root, the vertical direction became a relative one (varying from point to point on the terrestrial surface). It is well known that for a long time the Earth was thought to be the centre of the Universe. Within the framework of this model of the world, not all of the directions in space were considered identical, but only those that passed through the centre of the Earth. In other words, that model only provided for spatial isotropicity at one point, the Earth's centre; for any other point in space it was always possible to indicate physically different directions, for example the direction to the Earth's centre and one perpendicular to it. Transferring the centre of the Universe from the Earth to the Sun will, clearly, leave the situation unchanged, that is, some directions will remain singled out. And only a negation of all centres of the Universe squares with the idea of the isotropicity of space.

Symmetry in Time

One often hears that today's physics has nothing to do with the physics of earlier times. Even special terms are used to distinguish between them – "classical physics" and "modern physics". This terminology reflects the developments in the science of physics, which just as any other science does not and cannot be at a standstill.

The absolutely natural historical process of the development of physics in no way suggests that the laws of physics change over time. One of the most important symmetries of physical laws is their *constancy in time*, or rather their *invariance under a shift in time*.

The law of gravitation put forward by Newton describes the mutual attraction of bodies that does not change with the passage of time; the attraction had existed before Newton and it will exist in the centuries to come. The laws governing the behaviour of ideal gas, which were established in the 17th and 18th centuries, are widely used in modern science and engineering. No wonder that today's schoolchildren study, say, Archimedean principle, which was discovered in the 3rd century B. C. It will never occur to anybody that a TV set may go wrong because the physical laws governing the behaviour of the electron beam in electric and magnetic fields will change with time.

If the laws of nature changed with time, then each piece of research in physics would only have an "up-to-the-minute" significance. Each research worker would have to start from scratch, and there would be no continuity between generations of scientists, without which science cannot exist and progress. In a world without symmetry in time, the same causes would have certain effects today and other effects tomorrow.

The invariance of the laws of nature under shifts in time is normally called the *homogeneity of time.*

The Symmetry Under Mirror Reflection

Figure 88 presents the left-handed and right-handed molecules of butyl alcohol. One molecule is the looking-glass companion of the other; when reflected, each atom of the left-handed molecule coincides with the appropriate atom of the right-handed molecule, and vice versa. Suppose that the *atom-by-atom* mirror reflection was accomplished not with a single molecule but with a macroobject as a whole, for example with some physical device. As a result, we would obtain the looking-glass twin of the initial device that, up to the handedness, yields an exact replica not only of the *appearance* but also of the *internal* structure of the device, down to the atomic structure. If the original device contained right-handed helices or spirals, they will turn into left-handed ones; and right (left)-handed molecules or groups of molecules will turn into left (right)-handed ones. Such a mirror image may in principle exist as a real object (not only as a mental image).

Suppose then that we have the mirror image of a conventional clock, we will call it a "mirror" clock. It is quite obvious that such a clock will function in exactly the same way as the original clock. True, the hands of the "mirror" clock will turn in the opposite direction to the original clock, and so the face will look different as well. We might as well think of a "mirror" TV set, a "mirror" electric network, a "mirror" optical system, and so on. All of these must work in similar ways to the conventional devices, networks, systems, and so on. This means that we here have another symmetry of the laws of nature, the *invariance under mirror reflection.*

The following is a graphic illustration of such a symmetry. Suppose you sit in a cinema theatre, the back wall of which is replaced by a large flat mirror. If you turn about and watch the film in the mirror, you will notice nothing unusual physically. The events in the mirror screen will happen exactly as on the usual screen. True, on the mirror screen you will have difficulties reading inscriptions that might appear on it, and furthermore, a familiar landscape or familiar asymmetrical objects will appear unfamiliar. But the latter effect of nonrecognition associated with the replacement of "right" by "left" has nothing to do with the laws of physics.

Up until 1956 physicists treated mirror symmetry as being similar to those corresponding to homogeneity of space and time, isotropicity of space, invariance under the Lorentz transformations. Put another way, they held that mirror symmetry is inherent in all the laws of physics, without exception. In 1956 the American physicists Lee and Yang suggested that invariance under mirror reflection must not apply to the group of laws that describe the decay of elementary particles. This prediction was proved by a direct experiment in 1957. Much to their surprise, physicists found that mirror symmetry is inherent not in all physical laws and that in some phenomena nature exhibits *left-right*

asymmetry. We will take a closer look at that most interesting issue in Chapter 14. We will only note here that the discovery of mirror asymmetry in the decay of elementary particles seems to provide a clue to the striking fact of the asymmetry of living molecules (to be discussed in some detail in Chapter 7).

An Example of Asymmetry of Physical Laws

In order not to leave you with the impression that the laws of physics are invariant under any transformation, we will furnish an instructive example of transformations under which the laws of physics are *noninvariant*. Such an example is the transformations that involve changes in the spatial scale, or rather similarity transformations.

All the laws are noninvariant under similarity transformations. In other words, the geometrical similarity principle is, strictly speaking, inapplicable to the laws of physics.

True, the notion of similarity has struck deep roots in human mind. So it enjoys wide use in literature and the arts. Suffice it to remember Jonathan Swift, who sent his Gulliver first to Lilliput and then to the giants of Brobdingnag. The same idea lay at the foundation of conjectures ventured at the turn of the century that the atom is a solar system in the microscopic world.

It would seem that if we were to build a new set-up, each part of which would be several-fold larger (or smaller) than the appropriate part of the original set-up, the new installation would function in exactly the same way as the original one. It is not for nothing that aerodynamic and hydrodynamic structures are tentatively tested on scale models.

It is well known, however, for those dealing with the tests that models should not be scaled down too much. It was Galileo who first established that the laws of nature are not symmetrical with respect to changes in scale. He came to that conclusion while speculating about the strength of bones of animals as their size is increased. Similar reasoning is provided in Feynman's book *The Character of Physical Laws*. We can construct a toy gothic cathedral out of matches. Why then can we not construct a similar cathedral out of huge logs? The answer is: should we undertake the project the resultant structure would be so high and heavy that it would collapse. You might object that when comparing two things one should change everything that constitutes the system. Since the small cathedral is subject to the pull of the Earth, to be consistent we would have to subject the larger cathedral to the pull of the Earth increased appropriately, which would be even worse.

From the point of view of modern physics, the invariance of the laws of nature with respect to similarity transformations has a clear and comprehensive explanation – *atoms have a size whose order of magnitude is absolute, identical for the whole of the Universe*. Atomic size is determined by the universal physical constant – *Planck's constant* $\hbar$ ($\hbar = 1.05 \cdot 10^{-34}$ J·s); being given by the relationship $\hbar^2/me^2$, where m and e are the mass and charge of the electron, respectively (the minimum known mass at rest and the minimum known electric charge in nature).

This relationship yields 10^{-10} m for the linear size of the atom. It follows, by the way, that reducing the linear dimensions of some real installation having a volume of 0.1 m^3, say, billion-fold, we will then be left with only about one hundred atoms! Clearly, no workable apparatus can be manufactured out of such a small number of atoms.

A striking example of the asymmetry of physical laws with respect to scale changes is the fundamental fact that sufficiently strong scaling down puts the laws of classical mechanics, specifically Newton's laws, out of business. The *laws of* the microworld, *quantum mechanics*, take over.

10 Conservation Laws

There exist several basic laws of nature that have the mathematical form of conservation laws. A conservation law states that in a closed system some physical quantity, for example, the total momentum or energy, remains constant at all times.

J. Orear

An Unusual Adventure of Baron Münchhausen

You may have read the story of Münchhausen mired in a bog with his horse. The galant baron relates: "Now the whole of my horse's body has sunk into the stinking mud, now my head too began to sink into the bog with only the plait of my wig sticking above the water. What was to be done? We would have perished if it had not been for the prodigious strength of my hands. I am awfully strong. And so I took hold of my plait, jerked with all my might and easily pulled out of the bog both myself and my horse, whom I squeezed tightly with my legs like with pincers." The reader can easily catch the esteemed baron in the lie. Really, according to Newton's third law (action and reaction) the plait acts on the hand with a force equal in magnitude and opposite in direction to the force with which the baron's hand acts on the plait. And since the hand and the plait are parts of the *same* physical system (the baron), the resultant force exerted by the baron on himself will clearly be zero.

It is thus in principle impossible to lift oneself by the hair. In a more general context, this prohibition can be viewed as a consequence of the *momentum conservation law*. According to this law, the *momentum of a system will not change as a result of the interaction of the components of the system with one another*. As applied to our case, this means that the baron's momentum directed downwards into the mire cannot change as a result of the interaction of the baron's hand with his plait.

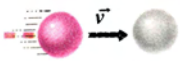
If the system is not subject to external influences (it is called *closed* system), then there are no reasons for the system's momentum to change. In that case, the total momentum does not change with time or, as it is conventionally put, it *is conserved*. The law of conservation of the *momentum* of a closed system is one of the three important conservation laws. The two other laws are the laws of conservation of *energy* and of *angular momentum*.

The Problem of Billiard Balls

The game of billiards provides a good opportunity to illustrate the action of conservation laws for *momentum* and *energy*. We will take a cue into our hands and try and push one of the billiard balls so that it rolls exactly along the line connecting the centre of this ball with the centre of another ball. This sort of collision is called a *head-on collision*. It is interesting that the first (striking) ball comes to rest at the moment of collision *no matter*

what its initial speed. The second ball starts moving along the line with a speed exactly equal to that of the first ball before the collision (Fig. 91).

This result can be computed using the laws of conservation of energy and momentum for the balls involved in the collision. Let v be the speed of the striking ball, the second ball being at rest before the collision. Further, let v_1 and v_2 be the speeds of the first and second balls, respectively, after the collision (at the moment we do not know that the first ball will come to rest after the collision). Billiard balls collide in an *elastic manner*, which means that no energy is lost in the collision. For an elastic collision the *law of conservation of energy* is



Before collision

$$\frac{mv^2}{2} = \frac{mv_1^2}{2} + \frac{mv_2^2}{2},$$

where m is the ball's mass. It follows that in the collision the total kinetic energy of the balls is conserved.

We will now turn to the *law of conservation of momentum.* Recall that the momentum of a body with mass m and speed $\vec{v}$ is $m\vec{v}$. Momentum is a *vector* quantity, and so we have to take into account its direction. In a head-on collision we have the only direction singled out physically (the red dash line in Fig. 91). Clearly, the momenta of the balls after the collision can only be directed along this direction. The *one-dimensional* nature of the problem enables us to consider, in the law of conservation of momentum, solely the numerical values of the momenta, that is, to write the law in its scalar form:

After collision

Fig. 91

$$mv = mv_1 + mv_2.$$

Combining both laws, we arrive at the set of equations in v_1 and v_2

$$\left.\begin{aligned} v_1^2 + v_2^2 &= v^2, \\ v_1 + v_2 &= v. \end{aligned}\right\}$$

Squaring both sides of the second equation gives

$$v_1^2 + 2v_1v_2 + v_2^2 = v^2.$$

From the first equation we obtain

$$2v_1v_2 = 0.$$

Physically, it is clear that $v_2 \neq 0$ (the second ball cannot remain at rest after the impact). We therefore conclude that $v_1 = 0$ and hence $v_2 = v$.

If billiard balls collide not in a central way (*off-centre* collision), then after the impact both balls will move apart in different directions. It is remarkable that in all cases the balls will move apart at a right angle, which can easily be tested in practice. But this result can be predicted from the laws of conservation of energy and momentum for colliding balls.

We will denote by $\vec{v}$ the speed of the striking ball, and by $\vec{v}_1$ and $\vec{v}_2$ the speeds of the balls moving off after the impact at an angle α (Fig. 92). Show that $\alpha = 90°$.

The *law of conservation of energy* will have the same form as for a central collision:

$$\frac{mv^2}{2} = \frac{mv_1^2}{2} + \frac{mv_2^2}{2},$$

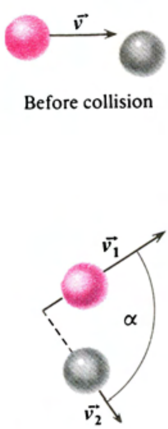


Fig. 92

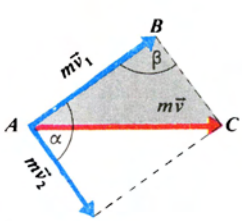


Fig. 93

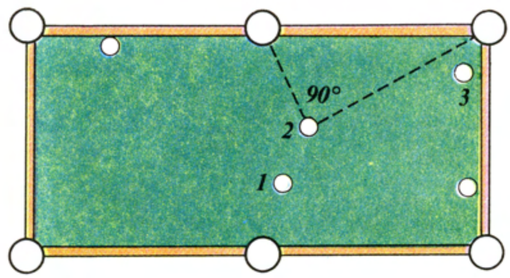


Fig. 94

but the *law of conservation of momentum* must now be written in a vector form:

$$m\vec{v} = m\vec{v}_1 + m\vec{v}_2.$$

Figure 93 presents vector $m\vec{v}$ as a sum of vectors $m\vec{v}_1$ and $m\vec{v}_2$. Let us consider the triangle ABC ($|AC| = mv$, $|AB| = mv_1$, $|BC| = mv_2$) and apply the *law of cosines*

$$|AC|^2 = |AB|^2 + |BC|^2 - 2|AB| \cdot |BC| \cos \beta,$$

where β is the angle between the sides AB and BC. This relationship can be written as

$$v^2 = v_1^2 + v_2^2 + 2v_1v_2 \cos \alpha$$

(considering that $\beta = 180° - \alpha$, and hence $\cos \beta = -\cos \alpha$).

This is the most convenient form of the law of conservation of momentum for the two colliding balls. Since from the law of conservation of energy $v^2 = v_1^2 + v_2^2$, we have $2v_1v_2 \cos \alpha = 0$. And since $v_1 \neq 0$ and $v_2 \neq 0$, then $\cos \alpha = 0$, that is $\alpha = 90°$.

Shown in Fig. 94 is a specific situation on a billiard table. The directions from ball *2* to the two pockets (see dash lines) form a straight angle. Therefore, it is better to drive ball *1* against ball *2*, not *3*, since a good stroke may send both ball *2* and ball *1* to the respective pockets.

On the Law of Conservation of Momentum

Let us take a closer look at the *law of conservation of momentum*. In our thought experiment we will replace the billiard table with a large flat

surface on which n balls move about colliding randomly with one another. We will suppose that in the general case the masses of the balls are different: $m_1, m_2, \ldots, m_n$. We will denote the speeds of the balls at time t as $\vec{v}_1(t)$, $\vec{v}_2(t)$, $\ldots$, $\vec{v}_n(t)$. We will then have the sum

$$m_1\vec{v}_1(t) + m_2\vec{v}_2(t) + \ldots + m_n\vec{v}_n(t),$$

which can be conveniently represented as

$$\sum_{i=1}^{n} m_i\vec{v}_i(t).$$

According to the law of conservation of momentum for a system of colliding balls, this sum must remain unchanged in time. Although individual summands here change in collisions, the sum as a whole remains a constant value (is conserved). The law of conservation of momentum can be written as

$$\sum_{i=1}^{n} m_i\vec{v}_i(t) = \vec{p},$$

where $\vec{p}$ is the total momentum of the balls. The vector $\vec{p}$ is constant, so that the collisions of balls with one another do not affect its direction and magnitude. The above equation implies that the *total momentum of some system* (in our case, a system of balls) *does not change when parts of the system interact with one another.*

In classical mechanics the law of conservation of momentum can be derived from *Newton's* third and second *laws*. The derivation is instructive, and so we will reproduce it here. Suppose that within a system of balls, two balls with masses m_1 and m_2 collide. Ball m_2 exerts a force $\vec{f}_1$ on ball m_1, and ball m_1 exerts a force $\vec{f}_2$ on ball m_2. According to Newton's third law,

$$\vec{f}_1 = -\vec{f}_2. \tag{1}$$

From Newton's second law, we have

$$\vec{f}_1 = m_1\vec{a}_1, \quad \vec{f}_2 = m_2\vec{a}_2,$$

where $\vec{a}_1$ and $\vec{a}_2$ are the accelerations of m_1 and m_2, respectively. Let Δt be the duration of the collision, and $\Delta\vec{v}_1$ and $\Delta\vec{v}_2$ the changes of the speeds of the balls during the collision. For small enough Δt we can assume that

$$\vec{a}_1 = \frac{\Delta\vec{v}_1}{\Delta t}, \quad \vec{a}_2 = \frac{\Delta\vec{v}_2}{\Delta t}.$$

As a result, equation (1) reads

$$m_1\frac{\Delta\vec{v}_1}{\Delta t} = -m_2\frac{\Delta\vec{v}_2}{\Delta t}.$$

We will now substitute $\Delta\vec{v}_1 = \vec{v}_1' - \vec{v}_1$, where $\vec{v}_1$ and $\vec{v}_1'$ are the speeds of the ball m_1 before and after the collision, respectively. In a similar way we obtain $\Delta\vec{v}_2 = \vec{v}_2' - \vec{v}_2$. Now equation (1) reads

$$m_1(\vec{v}_1' - \vec{v}_1) = -m_2(\vec{v}_2' - \vec{v}_2)$$

or

$$m_1\vec{v}_1' + m_2\vec{v}_2' = m_1\vec{v}_1 + m_2\vec{v}_2 .$$

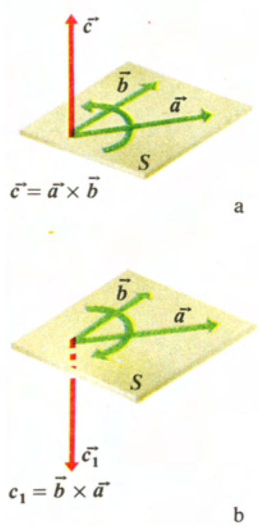


Fig. 95

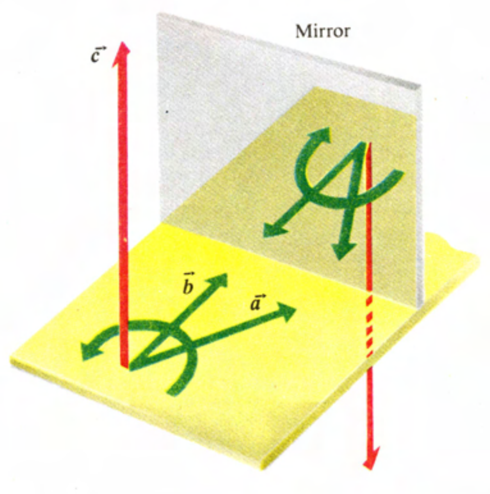


Fig. 97

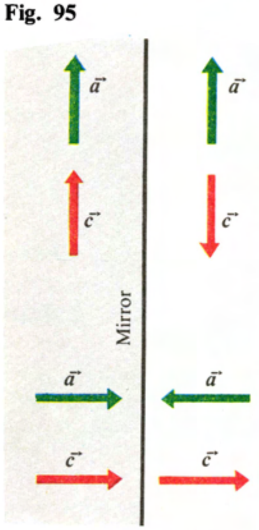


Fig. 96

The last relation implies that as a result of the collision of the balls their total momentum does not change. We thus arrive at the law of conservation of momentum.

The Vector Product of Two Vectors

Before we proceed to consider the law of conservation of angular momentum, we will take up a purely mathematical concept of the *vector product of two vectors*. Suppose that a vector $\vec{a}$ is to be multiplied in a *vector way* by a vector $\vec{b}$, or it is required to find a vector $\vec{c}$ that is the vector product of the initial vectors. In symbols we have $\vec{c} = (\vec{a} \times \vec{b})$. Let us translate one of the multipliers, say $\vec{b}$, parallel to itself so that both vectors ($\vec{a}$ and $\vec{b}$) have common origin. Denote by S the plane passing through $\vec{a}$ and $\vec{b}$ and by φ the angle between them. The magnitude of $\vec{c}$ will be given by

$$c = ab \sin\varphi,$$

and its direction is perpendicular to S. True, there are *two* directions perpendicular to the plane; these are mutually opposite. To have the

required direction for $\vec{c}$ we will turn the first multiplier ($\vec{a}$) to the second multiplier ($\vec{b}$) through the smaller angle. The direction of $\vec{c}$ will then obey the so-called *right-handed screw rule* (Fig. 95*a*). It is easily seen that transposing the multipliers changes the sign of their vector product: $(\vec{a} \times \vec{b}) = -(\vec{b} \times \vec{a})$ (compare Fig. 95*a* and *b*).

Note that the direction of $\vec{c}$ is an arbitrary notion, since it is associated with the convention to use the *right-handed screw*. Nothing forbids us in principle from starting off with the *left-handed screw*, not the *right-handed* one. Vectors, whose direction is not conditioned physically, but is arbitrarily related to the right (left)-handed screw, are called *axial* vectors. On the contrary, conventional vectors whose direction is defined physically are called *polar* vectors. The vector product of two polar vectors is an axial vector.

Notice that an *axial vector is reflected in a mirror in a different way from a conventional (polar) vector* (see Fig. 96, where $\vec{a}$ is a polar vector, $\vec{c}$ an axial vector). The handedness of the axial vector can be seen in Fig. 97, where the axial vector $\vec{c}$ is regarded as the vector product of the polar vector $\vec{a}$ and the polar vector $\vec{b}$. We can grasp the handedness of $\vec{c}$ by reflecting $\vec{a}$ and $\vec{b}$, and noticing the change in the *direction of rotation* from $\vec{a}$ to $\vec{b}$ (see circular arrow in Fig. 97).

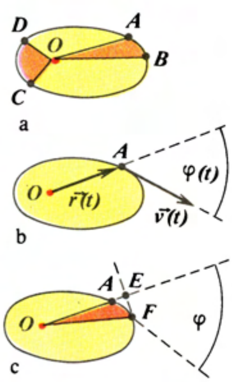


Fig. 98

Kepler's Second Law—the Law of Conservation of the Angular Momentum of a Planet

The second of Kepler's three laws describing planetary motion is known as the *law of areas*: an imaginary straight line connecting a planet with the Sun "sweeps" in the planet's orbital plane *equal areas in equal time periods.* Figure 98*a* shows the elliptical orbit of some planet, with the Sun being at focus *O*. The areas of the figures *AOB* and *COD* shaded in the figure are equal; therefore, by Kepler's second law, the planet covers segments *AB* and *CD* in equal times. Accordingly, the closer the planet approaches the Sun, the higher is its orbital velocity.

The law of areas discovered by Kepler corresponds to the *law of conservation of the orbital momentum* of the planet. We will clarify the concept of *angular momentum* with reference to Fig. 98*b*.

Suppose at a time t a planet lies at a point A on its orbit. We will draw from O to A a vector $\vec{r}(t)$ – the position vector of the planet at time t. The planet has a velocity $\vec{v}(t)$ and momentum $m\vec{v}(t)$ (where m is the planet's mass). The vectors $\vec{r}(t)$ and $\vec{v}(t)$ make an angle $\varphi(t)$. The *orbital angular momentum* of the planet, $\vec{M}$, is the vector product of the position vector $\vec{r}$ by the momentum vector $m\vec{v}$:

$$\vec{M} = (\vec{r}(t) \times m\vec{v}(t)).$$

Its magnitude is determined by the product $rmv \sin \varphi$, and its direction, according to the right-handed screw rule, is normal to the orbital plane (in Fig. 98 this is the direction into the page).

Here $\vec{r}$ and $m\vec{v}$ are conventional (*polar*) vectors, and $\vec{M}$ is an *axial* vector.

We will now show that *Kepler's law of areas corresponds to the law of conservation of the orbital angular momentum.* Let A and F be two close points on the orbit at which the planet lies at t and $t + \Delta t$, respectively (Fig. 98*c*). The time increment Δt is assumed to be small enough – such that the arc AF could be replaced by a segment and the orbital speeds of the planet at A and F could be considered essentially the same. The area of the shaded triangle AOF in Fig. 98*c* will be ΔS. Then from vertex F we will drop the perpendicular FE to the extension of AO. It is easily seen that

$$\Delta S = \frac{1}{2}|AO|\cdot|FE| = \frac{1}{2}|AO|\cdot|AF|\cdot\sin\varphi(t).$$

Since $|AO| = r(t)$ and $|AF| = v(t)\Delta t$, then

$$\Delta S/\Delta t = \frac{1}{2}r(t)\,v(t)\sin\varphi(t).$$

Thus,

$$\frac{\Delta S}{\Delta t} = \frac{M}{2m}.$$

The left-hand side of this equation includes the areas "swept out" by the position vector of the planet in a unit time, and on the right-hand side we have M, that is, the orbital angular momentum of the planet. According to Kepler's law, the quantity $\Delta S/\Delta t$ does not vary with time. It follows that M, too, must be independent of time.

True, Kepler's second law and the law of conservation of the orbital angular momentum of a planet are not completely equivalent. The law of conservation of momentum contains more information than Kepler's law of areas, since its content is the conservation of not only the *magnitude* but also of the *direction* of angular momentum in space. The conservation of the direction of angular momentum accounts for the fact that the orientation of the orbital plane of a planet is unchanged in space.

Conservation of the Intrinsic Angular Momentum of a Rotating Body

Apart from the *orbital* angular momentum, a planet also has an *intrinsic* angular momentum. Whereas orbital angular momentum is associated with the motion of a planet in its orbit, the intrinsic angular momentum arises as it rotates about its own axis. The intrinsic angular momentum of a planet is also conserved. Its direction makes an angle with the orbital plane, which does not change in the course of time. The conservation of intrinsic angular momentum is responsible for the constant alternation of night and day; the conservation of the direction of that momentum is responsible for the unchanged (for a given latitude) variation of the length of day during the different seasons.

The constancy of angular velocity and direction of rotation of

a gyroscope or common top is also related to the conservation of the intrinsic angular momentum of these bodies. When a figure skater pirouetting on ice extends his arms to come to a halt swiftly, he takes advantage of the law of conservation of intrinsic angular momentum. Extending the arms sidewards shifts some mass of the skater away from the axis of rotation, which by the law of conservation of angular momentum is compensated for by a reduction in the angular velocity of rotation. We will explain this using Fig. 99.

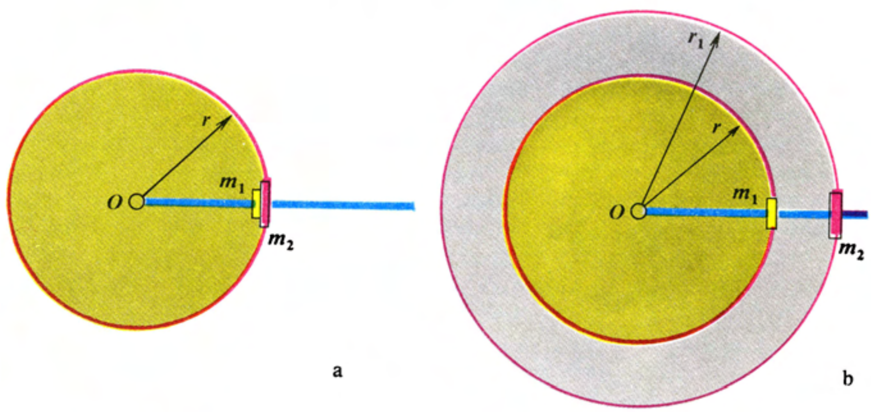


Fig. 99

For simplicity, in Fig. 99 instead of the skater we will consider a system of two masses m_1 and m_2 slid on a massless rod. The masses rotate on it about point O. In the initial situation both masses lie at the same distance r from O (see Fig. 99*a*). When the system is set in rotation, the *angular velocity* ω is related to the *orbital velocity* v by the relation $\omega = v/r$. Then the magnitude of the total angular momentum of the system will be

$$M = r(m_1 + m_2)v = (m_1 + m_2)\omega r^2.$$

We will further suppose that m_2 shifts to r_1 (see Fig. 99*b*). This more or less corresponds to the skater extending his arms. Now the angular momentum will be

$$M_1 = \omega_1 (m_1 r^2 + m_2 r_1^2).$$

Since $r_1 > r$ and $M_1 = M$ (the angular momentum is conserved), $\omega_1 < \omega$.

This law finds widespread use in engineering. So, this is the principle behind the gyroscope, a sort of top whose spin axis stays constant in space. Gyroscopes are used in gyrocompasses, automatic guidance devices, large gyros are used on some ships to achieve stability against rolling.

11 Symmetry and Conservation Laws

Today it is difficult to find a paper devoted to fundamental issues of physics that would not make mention of invariance principles and its author would not in his arguments draw on the concepts of the existence of relations between conservation laws and invariance principles.

E. Wiegner

The law of conservation of energy had been used in mechanics before Galileo. So in the late 15th century the great Leonardo da Vinci postulated the impossibility of perpetuum mobile. In his book *On True and False Science* he wrote: "Oh, seekers of perpetual motion, how many empty projects you have created in those searches." The laws of conservation of momentum and angular momentum were formulated later, in the 17-18th centuries. It was not, however, till the beginning of the 20th century that the laws assumed prominence. The attitude to them changed radically only after it had been discovered that *these laws were related to the principles of invariance.* Once this relation had been revealed, it became clear that conservation laws are predominant among other laws of nature.

The Relationship of Space and Time Symmetry to Conservation Laws

The relationship can be formulated as follows:

1. *The law of conservation of momentum is a consequence of the homogeneity of space*, or rather a consequence of the invariance of the laws of physics under translations in space. Momentum can thus be defined as a physical quantity whose conservation is a consequence of the above-mentioned symmetry of physical laws.

2. *The law of conservation of angular momentum is a consequence of the isotropy of space*, or rather a consequence of the invariance of the laws of physics under rotations in space. Angular momentum is a physical quantity that is conserved as a consequence of the above-mentioned symmetry of physical laws.

3. *The law of conservation of energy is a consequence of the homogeneity of time*, or rather a consequence of the invariance of the laws of physics under translations in time. Energy is a physical quantity that is conserved as a consequence of the above-mentioned symmetry of physical laws.

The three-dimensionality of space predetermines the vector nature of momentum and angular momentum, and so the laws of conservation of momentum and angular momentum are vector laws. The one-dimensionality of time predetermines the scalar nature of energy and the corresponding conservation law.

The relationship of conservation laws to space-time symmetry means that the passage of time or a translation and a rotation in space cannot cause a change in the physical state of the system. This will require that the system interact with other systems.

The Universal and Fundamental Nature of Conservation Laws

Such a way of looking at conservation laws may appear to be unusual to the uninitiated reader. Significantly, the *conservation laws can be obtained without applying the laws of motion directly from symmetry principles.* The derivation of the law of conservation of momentum from Newton's second and third laws, provided in Chapter 10, should in this connection be regarded as a *special* consequence.

It follows thus that the *region of applicability* of conservation laws is wider than that of the laws of motion. The laws of conservation of energy, momentum and angular momentum are used both in classical mechanics and in quantum mechanics, whereas Newton's laws of dynamics are out of place in quantum mechanics. Wiegner wrote: "For those who derive conservation laws from invariance principles, it is clear that the region of applicability of these laws lies beyond the framework of any special theories (gravitation, electromagnetism, etc.) that are essentially separated from one another in modern physics."

Clearly, the region of applicability of conservation laws must be as wide as that of appropriate invariances. This enables one to think of the laws of conservation of energy, momentum and angular momentum as *universal laws.*

The relationship of conservation laws to invariance principles also suggests that any violation of these laws, should it occur, would point to a violation of appropriate invariance principles. So far no experimental evidence is available to indicate that the laws of nature may turn out to be noninvariant under translation in time, and under translation and rotation in space. This circumstance, together with the aforementioned property of universality, makes the laws of conservation of energy, momentum, and angular momentum really *fundamental laws.*

It follows from the fundamental nature of conservation laws that we should select as the *conserved quantities fundamental physical quantities*: energy, momentum, and angular momentum. Note that in classical mechanics these quantities appear as functions of the velocity and coordinates of a body. So the energy, momentum, and angular momentum of a billiard ball can be represented in the form

$$E = \frac{mv^2}{2}, \quad \vec{p} = m\vec{v}, \quad \vec{M} = (\vec{r} \times m\vec{v}), \tag{1}$$

whence it follows in particular that

$$E = \frac{p^2}{2m}, \quad \vec{M} = (\vec{r} \times \vec{p}). \tag{2}$$

It might appear that we can conclude from expressions of type (1) that the role of fundamental quantities is played by velocity and coordinates. But if we go over from the classical mechanics of billiard balls to the quantum mechanics of microobjects, the very concept of the velocity of an object will then become unsuitable and expressions (1) become pointless. At the same time the conserved quantities (E, $\vec{p}$, $\vec{M}$) retain their

meaning both in classical and quantum mechanics. It is important that in quantum mechanics they are, generally speaking, not expressible in terms of each other; the second of (2) does not hold in the microworld, since a microobject has no states in which the values of momentum and coordinates are specified simultaneously. As regards the first of (2), it is valid only for the free motion of a microobject. For a bound microobject (for instance, an atomic electron) the energy is quantized, with the result that for each energy level we cannot indicate a definite value of momentum. Quantum mechanics, thus, actually makes it possible to bring out the fundamental nature of conserved quantities, their independence in full conformity with the fundamentality and independence of the corresponding types of symmetry of the laws of physics.

The universality of conservation laws suggests that the conserved physical quantities are used in *various branches of physics*. They can be described by various expressions into which enter physical quantities characteristic of a given field. Consider momentum, for example. We will write out the four expressions for the momentum

$$\vec{p} = m\vec{v};$$

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}};$$

$$\vec{p} = \frac{1}{c^2}(\vec{E} \times \vec{H});$$

$$\vec{p} = \frac{\hbar\omega}{c}\vec{n}.$$

The first expression describes the momentum of a body with mass m and speed $\vec{v}$ in classical mechanics; the second one, in the theory of relativity. The third expression is the momentum of a unit volume of electromagnetic field in terms of the vectors of the strengths of electric and magnetic fields. The fourth expression gives the momentum of a photon in terms of its cyclic frequency ω and unit vector $\vec{n}$ in the direction of the motion of the photon; here $\hbar$ is Planck's constant. These four formulas are a good illustration of the universal nature of the concept of "momentum". As for the quantities $\vec{v}$, $\vec{E}$, $\vec{H}$, and ω, they are all applicable in respective disciplines of physics only.

It is worth noting here that proceeding from Newton's laws of dynamics we can obtain the law of conservation of momentum solely for the special case where the momentum is given by the expression $\vec{p} = m\vec{v}$; in other cases Newton's laws are clearly not applicable. If we start off with invariance principles, we can derive the law of conservation of momentum regardless of the expression describing the momentum in one case or another.

The Practical Value of Conservation Laws

In the introductory talk about symmetry, it was noted that symmetry is some general entity inherent in a wide variety of objects (phenomena), whereas asymmetry brings out some individual characteristics of a specific object or phenomenon. Permeating all spheres of physics and all specific situations, *conservation laws express those general aspects of all situations, which is eventually related to the appropriate principles of symmetry*. These laws "ignore" the particularity of a given situation, they "ignore" the particular mechanisms of interplay, their region of applicability *lies beyond the framework of specific theories*. It is the general, all-embracing character of conservation laws, which do not require an analysis of phenomena, that is responsible for the striking simplicity of these laws and unconditioned reliability of results derived from them.

It is worth mentioning that not infrequently the interaction mechanisms (details of a phenomenon) are unknown or known in only a fairly approximate way. In many cases the inclusion of a host of details in a problem overly complicates the mathematical side of the problem. Against the background of these difficulties, the simple and elegant conservation laws appear quite attractive. When handling a phenomenon, a physicist, above all, applies conservation laws and only then proceeds to examine details. Many phenomena to date have been studied on the level of conservation laws only.

So in the billiard problem (see Chapter 10) we neglected the mechanism of collision of balls. Otherwise, we would have had to delve into the details of elasticity theory. In this particular case, it turned out to be sufficient to utilize the laws of conservation of energy and momentum alone.

The Example of the Compton Effect

We will illustrate the practical value of conservation laws referring to an effect discovered in 1923 by the British physicist A. Compton (the *Compton effect*). The gist of the effect is as follows: when a monochromatic beam of X-rays is scattered at electrons that enter the composition of the target substance, the wavelength of the radiation is increased by $\Delta\lambda$, which is found using the relation (the *Compton formula*)

$$\Delta\lambda = \frac{4\pi\hbar}{mc}\sin^2\frac{\varphi}{2},$$

where φ is the scattering angle, $\hbar$ Planck's constant, m electron rest mass, c the velocity of light in a vacuum.

To derive the Compton formula we need not know the exact mechanism of the interaction of X-ray quanta (photons) with the electrons; *it is sufficient to apply the laws of conservation of energy and momentum for a collision of a photon with an electron*. In a sense, we again have to turn to the billiard problem, although neither the electron nor the photon resemble a billiard ball, of course. Before a collision, the ball of an electron can be treated as being at rest as compared with the incident ball of a photon. The energy of the photon is E, its momentum is $\vec{p}$. Assuming

the collision to be off-centre, after it, the photon will be scattered at an angle φ, its energy being E_1 and momentum $\vec{p}_1$; the electron will be scattered at an angle θ, its energy being E_e and momentum $\vec{p}_e$ (Fig. 100). *The law of conservation of energy has the form*

$$E = E_1 + E_e,$$

and the *law of conservation of momentum* (Fig. 101)

$$\vec{p} = \vec{p}_1 + \vec{p}_e,$$

or in the components,

$$\left.\begin{aligned} p &= p_1 \cos\varphi + p_e \cos\theta, \\ 0 &= p_1 \sin\varphi - p_e \sin\theta. \end{aligned}\right\}$$

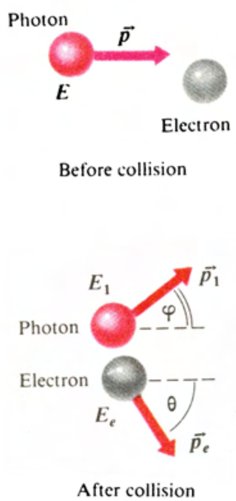


Fig. 100

The form of these equations is *quite general*, they are applicable also for billiard balls. To take account of the specific feature of the photon-electron problem it is sufficient to express the photon energy as

$$E = \hbar\omega,$$

and its momentum, using the formula discussed earlier in the book, as

$$\vec{p} = \frac{\hbar\omega}{c}\vec{n}.$$

As regards the electron, its energy can be given in terms of momentum using the conventional relation

$$E_e = \frac{p_e^2}{2m}$$

that is equally valid both for the billiard balls and for the free electron (provided that its velocity is well under the velocity of light).

It follows that the conservation laws can be written as

$$\hbar\omega = \hbar\omega_1 + \frac{p_e^2}{2m},$$

$$\frac{\hbar\omega}{c} = \frac{\hbar\omega_1}{c}\cos\varphi + p_e\cos\theta,$$

$$0 = \frac{\hbar\omega_1}{c}\sin\varphi - p_e\sin\theta.$$

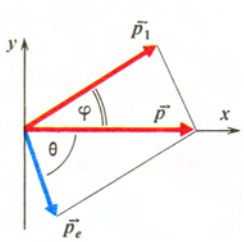


Fig. 101

It only remains to do some algebra. Introducing the notation $\Delta\omega = \omega - \omega_1$ ($\Delta\omega$ is generally small, that is, $\Delta\omega \ll \omega$), we can write the above set of equations in the form

$$\hbar\Delta\omega = p_e^2/2m,$$

$$\hbar\omega - \hbar(\omega - \Delta\omega)\cos\varphi = p_e c\cos\theta,$$

$$\hbar(\omega - \Delta\omega)\sin\varphi = p_e c\sin\theta.$$

Squaring the second and third equations and adding them together, we get rid of θ:

$$\hbar^2\omega^2 + \hbar^2(\omega - \Delta\omega)^2 - 2\hbar^2\omega(\omega - \Delta\omega)\cos\varphi = p_e^2 c^2,$$

or, using the first equation (the law of conservation of energy),

$$\omega^2 + (\omega - \Delta\omega)^2 - 2\omega(\omega - \Delta\omega)\cos\varphi = 2mc^2\Delta\omega/\hbar.$$

We will now divide both sides of this by ω^2 and open the brackets. Ignoring the small term $(\Delta\omega/\omega)^2$ gives

$$1 - \frac{\Delta\omega}{\omega} - \left(1 - \frac{\Delta\omega}{\omega}\right)\cos\varphi = \frac{mc^2}{\hbar\omega}\cdot\frac{\Delta\omega}{\omega}$$

or

$$2\sin^2\frac{\varphi}{2}\left(1 - \frac{\Delta\omega}{\omega}\right) = \frac{mc^2}{\hbar\omega}\cdot\frac{\Delta\omega}{\omega}.$$

The wavelength λ is related to the cyclic frequency ω by the expression $\lambda = 2\pi c/\omega$. It is easily seen that

$$\Delta\lambda = \lambda_1 - \lambda = \frac{2\pi c}{\omega_1} - \frac{2\pi c}{\omega} = 2\pi c\frac{\omega - \omega_1}{\omega\omega_1} = 2\pi c\frac{\Delta\omega/\omega^2}{1 - \Delta\omega/\omega}.$$

Hence

$$2\sin^2\frac{\varphi}{2} = \frac{mc}{2\pi\hbar}\Delta\lambda.$$

The Compton formula follows from this immediately.

We have thus seen that by approaching the Compton effect at the level of conservation laws, we can derive the Compton formula, that is, to find the variation of the wavelength increment with the scattering angle for the photon.

Before we leave this section, it should be stressed that it would be wrong to overestimate the role of conservation laws. To be sure, they are not a panacea for all aches and pains of the physicist. So to determine the probability for a photon to scatter at a given angle, it is necessary, apart from conservation laws, to take into account the specific mechanism of the interaction of photons and electrons. In that case, the parallel with billiard balls does not work any more.

Conservation Laws as Prohibiting Rules

It is well known that conservation laws are often formulated as *rules (laws) of prohibition.* So the law of conservation of energy is, in essence, the law that prohibits perpetual motion. In a sense the law of conservation of momentum is the law that prohibits lifting oneself by the hair. The law of conservation of angular momentum prohibits, for example, a planet from leaving its orbit and changing the angle at which its spin axis is inclined to the orbital plane.

There are two reasons to approach conservation laws as laws of prohibition. Above all, we notice that *symmetry, which introduces a measure of orderliness, at all times tends to reduce the number of possible versions*. It has already been stressed in Chapter 4 that symmetry dramatically limits the diversity of structures that may be encountered in nature. We can now complement that statement by stressing that symmetry limits the diversity of *not only structures* but also of the *forms of behaviour of physical systems*. Through conservation laws (or rather prohibition laws) symmetry covers all the conceivable forms of the behaviour of a system and at times it nearly uniquely predetermines a certain behaviour of a system.

Remember the billiard balls. No matter how strongly you push a ball to a head-on collision it is bound to come to rest, since by energy and momentum conservation it may not move after the impact. Also of interest is the case of off-centre collision: the conservation laws decree that after the collision the balls only move off at right angle to each other.

a

b

Fig. 102

In his book *The World of Elementary Particles* the American physicist K. Ford writes: "The older view of a fundamental law of nature was that it must be a law of *permission*. It defined what *can* (and must) happen in natural phenomena. According to the new view, the more fundamental law is a law of *prohibition*. It defines what *cannot* happen. A conservation law is, in effect, a law of prohibition. It prohibits any phenomenon that would change the conserved quantity."

It would seem that *prohibitory* rules represent just a simplified version of *guiding* rules. In actual fact, this is not so. Take road signs, for example. Suppose the sign shown in Fig. 102*a* is put up before a crossing. This sign is guiding: it explicitly directs the traffic – only forward. If then the sign is as shown in Fig. 102*b* (prohibitory sign), the traffic does not have an unequivocal direction – it may either continue forward or turn to the left; it is only prohibited to turn to the right. The meaning of the above quotation from Ford is that the conservation laws are to be likened to prohibitory laws, and not guiding laws.

There is one more important reason for which conservation laws are to be regarded as laws of prohibition. The fact is that in the world of elementary particles (where, additionally, there are a wide variety of other conservation laws), conservation laws are derived as rules that prohibit the phenomena that are never observed in experiment. Suppose experimental data definitely indicate that some transmutations of particles never occur, or in other words are prohibited. This circumstance may be used as a basis for formulating *some conservation law*. But it is not always that the invariance principle underlying the law is revealed; in such cases a conservation law only appears as a law of prohibition. We will be looking at such examples later in the book.

12 The World of Elementary Particles

The elementary particles are not just interesting scientific curiosities. They represent the deepest-lying substructure of matter to which man has been able to probe; consequently, they provide one of the most challenging problems on the current frontiers of science.

K. Ford

Elementary particles are the frontier of modern physics which corresponds to the most fundamental level of probing into the physical picture of the world. Naturally, it is here that the most important regularities show up, which, in the final analysis, control the structure of matter and the character of physical processes. It is of principal importance, therefore, to gain an insight into the aspects related to conservation and invariance in the world of elementary particles.

But before we proceed to discuss these issues, we will have to take a look at the currently known elementary particles.

Some Features of Particles

We will single out three quantities characterizing particles – *mass, electric charge,* and *spin.* The ensemble of characteristics will be extended markedly later. In addition, it will include lifetime, specific charges (electronic, muonic, and baryonic), isospin, strangeness, and so on.

By mass we will understand the *rest mass* of a particle, that is, the mass in the frame of reference connected with the particle itself. The smallest mass is possessed by the electron ($m = 9.1 \cdot 10^{-28}$ g); therefore, the mass of other particles is often expressed in electron masses. Mass is also expressed in energy units MeV (megaelectronvolts). The use of energy units for mass is based on the well-known relationship due to Einstein: $E = mc^2$. In terms of energy units, the electron mass is 0.511 MeV.

The *electric charge* of particles is denoted by numerals: 0, $+1$, -1. In the first case, there is no charge (the particle is neutral). In the second case, the charge is equal to that of an electron, but unlike the electron, it is positive. In the third case, the charge coincides with the electron charge both in magnitude and in sign. Note that the electric charge of charged particles is exactly equal to the electron charge, that is, $1.6 \cdot 10^{-19}$ C.

The *spin* of a particle is the specific angular momentum of a particle which can be called the *intrinsic angular momentum* since it is not related to motions of the particle in space; it is indestructible, its magnitude is independent of external conditions. This angular momentum can arbitrarily be associated with the rotation of the particle about its own axis. An analogue of spin may be the intrinsic angular momentum of a planet or gyroscope as discussed in Chapter 10. The squared spin is given by the expression $\hbar^2 s(s+1)$, where $\hbar$ is Planck's constant, s is a number characterizing this particle, which is normally referred to as its

spin (in the latter case spin is measured in units of $\hbar$). Like any angular momentum, spin is a vector quantity. This vector is quite specific, however, since its *projection* in a given fixed direction only takes on *discrete* values (is quantized): $\hbar s, \hbar(s-1), \ldots, -\hbar s$. The total number of spin projections is $2s+1$. In this connection, it is said that a particle with spin s may be in one of $2s+1$ *spin states*.

For many elementary particles, specifically for the electron, spin is 1/2. These particles have two spin states, one for each of the opposite spin directions.

Note that all the particles of this type (for example, all electrons) have exactly the same mass, charge, and spin. It is in principle impossible for the mass of one electron to differ from that of another one by, say, 0.001 per cent. The values of the mass, electric charge, and spin of the electron are, according to the evidence available, the lowest values of these quantities ever encountered in nature, save for the cases where a particle does not have a rest mass, charge, or spin.

The Zoo of Elementary Particles

The particles in the zoo are generally classed according to their mass, charge, and spin. They form three families:

The *first family* is the smallest – it consists only of one particle. It is the *photon*, the quantum of electromagnetic radiation (symbol γ). The rest mass and electric charge of the photon are zero, $s=1$. Note that according to the theory of relativity, any particle with zero rest mass cannot have an electric charge and in any inertial frame of reference it travels with the same velocity – the velocity of light in empty space. The photon is an example of such a particle.

The *second family* consists of particles called *leptons*. Up until 1975 four leptons were known: *electron* (e^-), *electron neutrino* (ν_e), *muon* (μ^-), and *muon neutrino* (ν_μ). We have already discussed the electron. The muon has a mass of $207m$, its electric charge is negative, $s = 1/2$. Both neutrinos are indistinguishable in terms of the three characteristics used here (no wonder that for a long time it was believed that there only exists one type of neutrino in nature). Like the photon, both neutrinos have neither rest mass nor electric charge. Unlike the photon, however, the spin of the neutrino is 1/2, like that of any lepton.

In 1975 a fifth lepton, *tauon* (τ^-), was discovered. It appeared to be an ultra-heavy particle: its mass is about $3500m$. The electric charge of the tauon is negative. Physicists have good reasons to think that a tauon must have a companion particle, a *tauon neutrino* (ν_τ). Counting the third type of neutrino, the number of leptons becomes six.

The *third family* consists of particles called *hadrons* (from the Greek for "large", "massive"). Hadrons are numerous: several hundred of them are known.

The hadron family divides into two subfamilies: the *mesons* and the *baryons*. The mesons either have no spin or have integer spins, whereas the baryons have half-integral spins. Among the hadrons (both mesons and baryons) there are many particles that decay quickly, their lifetime is

only 10^{-22}-10^{-23} s; these particles are called *resonances*. If we don't include the resonances, then up until 1974 physicists have identified 14 hadrons, among them five mesons and nine baryons.

The five mesons include two *pions* (the neutral pion π^0 and the positively charged pion π^+), the *kaons* (the positively charged kaon K^+ and the neutral kaon K^0), the neutral *eta meson* (η^0). All of these mesons are spinless particles ($s = 0$). Their masses are as follows: $\pi^0 - 264\,m$, $\pi^+ - 273\,m$, $K^+ - 966\,m$, $K^0 - 974\,m$, $\eta^0 - 1074\,m$.

The nine baryons include the *nucleons* (the proton p and the neutron n), the neutral *lambda hyperon* Λ^0, the *sigma hyperons* (the neutral one Σ^0 and the charged ones Σ^+ and Σ^-), the *xi hyperons* (the neutral one Ξ^0 and the negatively charged one Ξ^-), the negatively charged *omega hyperon* Ω^-. The omega hyperon has the spin 3/2; the other baryons have $s = 1/2$. The masses of the above baryons are as follows: $p - 1836.1\,m$, $n - 1838.6\,m$, $\Lambda^0 - 2183\,m$, $\Sigma^+ - 2328\,m$, $\Sigma^0 - 2334\,m$, $\Sigma^- - 2343\,m$, $\Xi^0 - 2573\,m$, $\Xi^- - 2586\,m$, $\Omega^- - 3273\,m$.

Note that the particles of one group in the meson or baryon subfamily have similar masses. So with pions they only differ by 3 per cent, with kaons, by 0.7 per cent, with nucleons, only by 0.14 per cent. The members of a group mainly differ by the electric charge. In this connection the mesons or baryons of one group can be viewed as one particle characterized by several *charge states*. With this approach the baryons Σ^+, Σ^0, Σ^- are one particle (sigma hyperon) that may be in three different charge states. (The small difference in masses of the charge components is due to the difference in the sign of the electric charge.) Pion, kaon, nucleon, and xi hyperon have two charge states each. Figure 103 presents all the above elementary particles, with the exception of those with zero rest mass. The vertical axis in the figure shows the mass; the horizontal one, the electric charge. The yellow rectangles combine those mesons or baryons that may be treated as different charge states of a particle.

Particles and Antiparticles

The diagram of Fig. 103 contains 17 particles. As a matter of fact, the number of particles to be considered is to be doubled. The fact is that each particle, with the exception of the photon, the neutral pion, and the eta meson, corresponds to an *antiparticle*. So for the electron we have its antiparticle, the positron (e^+); there are two antineutrinos, the electron one ($\bar{\nu}_e$) and the muon one ($\bar{\nu}_\mu$), and so on. The photon, the neutral pion, and the eta meson do not have antiparticles. We can say that each of these particles is identical to its antiparticle. Such particles are referred to as *truly neutral* ones.

Including the antiparticles, the number of particles at hand becomes 37. The diagram in Fig. 104 contains all of these particles, save for zero mass ones (photon, two neutrinos and two antineutrinos). The particles in the figure are represented by the blue colour, the antiparticles by the red colour, and the truly neutral particles by the green colour. The diagram is far from comprehensive: it does not include the short-lived

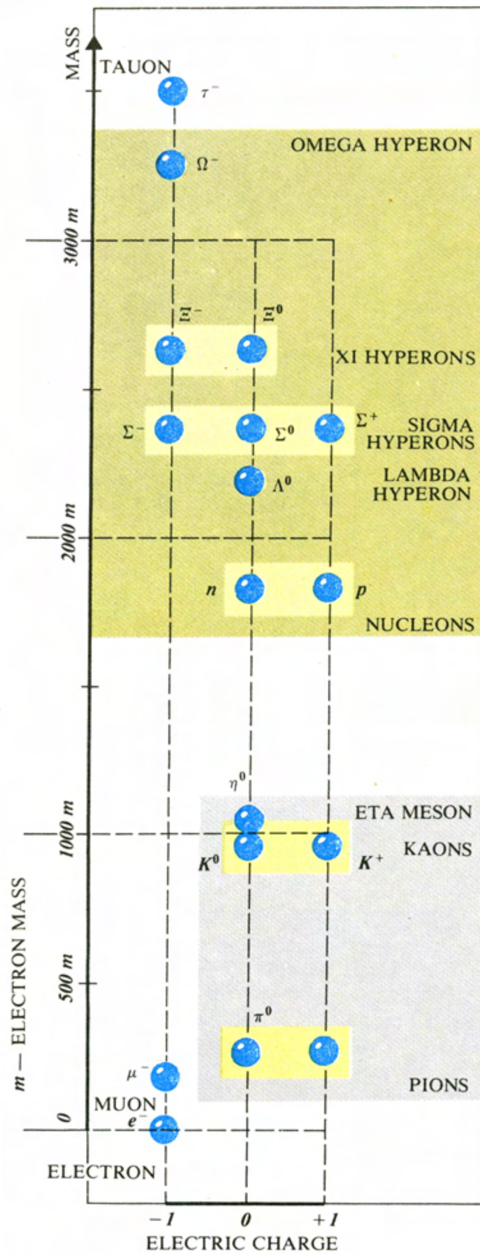


Fig. 103

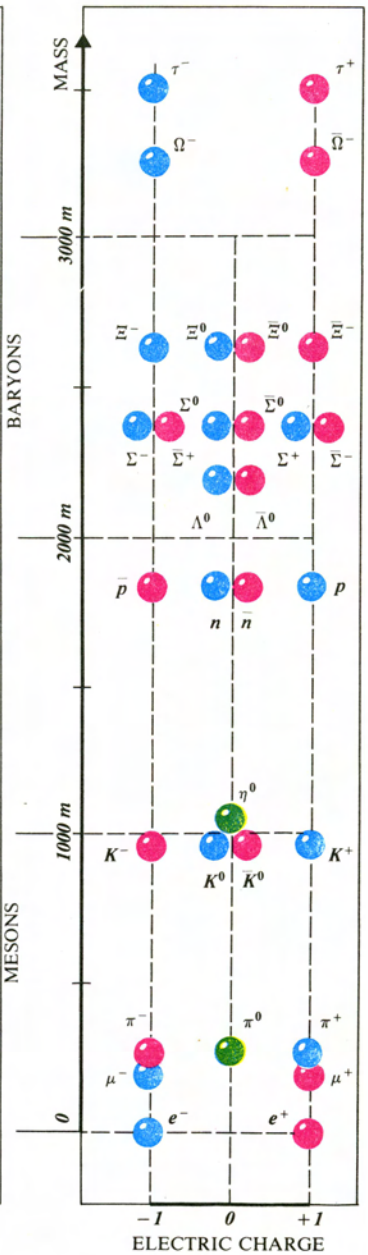


Fig. 104

hadrons (resonances) and the hadrons discovered since 1974 (the so-called charmed hadrons).

Particles and their antiparticles have equal spins and masses, but the signs of their electric charges are opposite. So, unlike the proton, the antiproton is negative. But what if a particle is not charged? What is the difference, say, between a neutrino and antineutrino, or a neutron and antineutron? We will answer the question in more detail later in the book, for the moment we will only note that elementary particles are characterized not only by electric charge, but also by a number of other charges, which will be considered in Chapter 13. *Particles and antiparticles have opposite charges.*

It is seen in Fig. 104 that antiparticles are normally denoted by a bar over the symbol for a particle. So $\overline{\Lambda^0}$ is the antiparticle for the lambda hyperon Λ^0, or antilambda hyperon. Since the charges of a particle and its antiparticle have opposite signs, in a number of cases (electron, muon, charged mesons) no bar is used. The electron is denoted by e^-; the antielectron (positron) by e^+. So e^+, μ^+, π^-, K^- are symbols for antiparticles, whereas e^-, μ^-, π^+, K^+ are symbols for particles. But in the case of the sigma hyperon Σ^+, we cannot denote the antiparticle by Σ^-, because there exists the sigma hyperon Σ^-. It is absolutely necessary here to use for the antiparticle the symbol $\overline{\Sigma^+}$. It should be remembered that the antisigma hyperon $\overline{\Sigma^+}$ has a negative electric charge.

But whatever the similarity between a particle and its antiparticle, they differ in a fundamental way: particle and antiparticle annihilate in merging with the result that several mesons or photons are formed. So Σ^+ and $\overline{\Sigma^+}$ are mutually destroyed, whereas the encounter of Σ^+ and Σ^- is quite peaceful. The mutual destruction process (or in the language of physics, *annihilation*) may be an indication of a meeting of a neutron and antineutron, and not of two neutrons.

Let us take some examples of events (reactions) of annihilation of particle and antiparticle:

$$e^- + e^+ \rightarrow \gamma + \gamma,$$

$$p + \bar{p} \rightarrow \pi^+ + \pi^- + \pi^0,$$

$$p + \bar{p} \rightarrow \pi^+ + \pi^+ + \pi^- + \pi^-,$$

$$p + \bar{p} \rightarrow \pi^+ + \pi^+ + \pi^- + \pi^- + \pi^0.$$

The first reaction is the annihilation of an electron and a positron yielding two photons. The other three reactions are the various cases of annihilation of a proton and an antiproton.

Particles, Antiparticles and Symmetry

The very fact that antiparticles exist is closely connected with symmetry. Without antiparticles the equations of theoretical physics describing the various types of elementary particles would be noninvariant under the Lorentz transformations, that is, under transfer from one inertial frame to another. Put another way, *the existence of antiparticles in addition to*

particles is related to the invariance of the laws of physics under transfer from one inertial frame of reference to another. This, unfortunately, lies beyond the scope of this book.

We live in a world of particles, antiparticles being fairly rare visitors. One can, however, imagine another world, the one built of antiparticles. We will call it the *antiworld.* Antimatter in the antiworld consists of antiatoms and antimolecules. So the antiatom of hydrogen has the nucleus of an antiproton, with a positron orbiting in its field. In the antiatom of helium, two positrons orbit in the nuclear field, the nucleus consisting of two antiprotons and two antineutrons.

An encounter with the antiworld is a favourite theme of science fiction. After many years of flight at a speed near the speed of light a spaceship arrives at an unknown planet. The ship starts orbiting the planet and the astronauts start studying the planet's surface, using a most advanced optical apparatus that enables them to discern things up to one metre across from a height of several hundred kilometres. They find that the planet is inhabited and, moreover, the intelligent creatures look very much like people. At the same time, the astronauts endeavour to contact the dwellers of the planet. After having spent, say, a fortnight in orbit, the astronauts become confident that they have found a similar civilization which strikingly resembles their own terrestrial one. Meanwhile, the space travellers and the extraterrestrials reach a measure of mutual understanding. The astronauts are kindly invited to land on the planet at a specified region. The terrestrial ship starts its engines and begins descending, approaching the lower atmosphere of the hospitable planet. And now a catastrophe happens: the shocked extraterrestrials see, high up in the sky where the spaceship was just visible, a blinding flash...

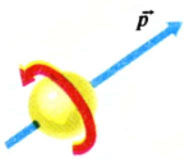
Today's science fiction reader or fan has already twigged that the hospitable planet is a piece of antiworld. Upon entry into the antimatter atmosphere, the terrestrial spacecraft was immediately destroyed due to the annihilation of the particles of the ship with the antiparticles of the atmosphere.

In this science fiction story there is one moment that is worth examining. It is not surprising that the astronauts made the fatal mistake: the antiworld of the planet must have appeared as natural as the world of their Earth. The exchange of information by radio between the ship's crew and the planet could also give no grounds to suspect that they belonged to opposite worlds. *World and antiworld are absolutely symmetrical; the laws of nature are invariant under replacement of all the particles by their respective antiparticles.* This invariance of the laws of physics is normally referred to as *charge invariance* (or *C-invariance*). It is not surprising then that when viewing from a distance a world built of antiprotons, antineutrons, positrons, etc., the crew mistook it for a conventional world, the one built of protons, neutrons, electrons, etc. By the way, they contacted that world by the agency of photons, that is, truly neutral particles, which are the same for both the conventional world and the antiworld.

But the charge invariance of the laws of nature is not a completely rigorous invariance: there are processes in which it does not hold. And so, in principle, the space travellers could have made some experiments on board their ship to clarify the nature of the world before them to determine whether it was a conventional world or an antiworld. We are going to take a closer look at this later in the book.

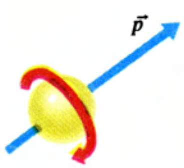
The symmetry of the physical properties of the world and the antiworld is combined with the distinct asymmetry of the spatial distribution of matter and antimatter. It has been estimated that in our Galaxy, for each antiparticle there are more than 10^{17} particles. It follows, in particular, that the probability for a spacecraft to hit upon a star system made from antimatter is infinitesimal.

This asymmetry has not yet been explained. It is to be assumed that either the Universe as a whole is charge asymmetric or matter and antimatter have spatially separated into isolated regions, which interact with one another exceedingly weakly. Both assumptions pose a variety of fundamental questions which, so far, scientists have not been able to answer.



Neutrino

a



Antineutrino

b

Fig. 105

Neutrino and Antineutrino (Left and Right Helices in the World of Elementary Particles)

The spin of a neutrino is always aligned *against* (antiparallel to) the neutrino's momentum. This means that if we visualize the spin through the rotation about the neutrino's axis, the latter is always aligned with the direction of motion of the particle. If we follow a neutrino flying off, it will spin counterclockwise (Fig. 105*a*). The antineutrino's rotation axis is also parallel to the direction of motion. Unlike neutrino, however, the receding antineutrino rotates clockwise (Fig. 105*b*). In other words, *the neutrino can be compared to a left-handed helix and the antineutrino to a right-handed helix.* This holds good both for the electronic and for the muonic neutrino (antineutrino).

It might appear that this model of neutrino (antineutrino) contradicts the principle of invariance with respect to transfers from one inertial frame of reference to another one. Suppose that a left-handed helix is flying past us. In other words, it is a left-handed helix in the laboratory frame of reference. Suppose further that we try to catch up with the neutrino in a spacecraft travelling at a velocity higher than the neutrino's velocity. Then in the frame associated with the ship the neutrino will now move not away from us but toward us, the counterclockwise rotation of the neutrino will not change in the process. This means that in the ship's frame of reference the neutrino will be a right-handed helix, not a left-handed one.

This reasoning is invalid, however. The fact is that in any inertial frame, a neutrino travels at the velocity of light (recall that the neutrino's rest mass is zero); therefore the ship cannot have a velocity higher than that of the neutrino. And so the neutrino remains a left-handed helix in any frame of reference.

Note. The attentive reader may have noticed that such a model of neutrino (antineutrino) does not agree with the invariance of the laws of physics under mirror reflection. A mirror reflection turns the left-handed helix into a right-handed one, and so a neutrino must turn into an antineutrino. The reader is absolutely right. We will be looking at this issue in more detail in Chapter 14.

The Instability of Particles

One of the most important features of elementary particles is that most of them are *unstable*. This means that particles decay spontaneously, without any induction, yielding other particles. Exceptions are the photon, neutrino, electron, and proton (with respective antiparticles); these subatomic particles do not decay, they are stable.

A free neutron, for instance, decays spontaneously to produce three stable particles, namely, a proton, an electron and an electron antineutrino*:

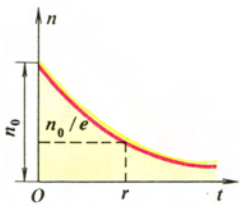


Fig. 106

$$n \to p + e^- + \bar{\nu}_e.$$

A positively charged pion decays to yield an antimuon and a muon neutrino

$$\pi^+ \to \mu^+ + \nu_\mu;$$

the resultant antimuon decays, in turn, to give a positron, a muon antineutrino, and an electron neutrino

$$\mu^+ \to e^+ + \bar{\nu}_\mu + \nu_e.$$

A neutral pion disintegrates into two photons

$$\pi^0 \to \gamma + \gamma.$$

Table 1 at the end of the chapter tabulates the principal decay schemes for unstable particles.

The decay of an elementary particle is a phenomenon that is in need of some explanation. To begin with, we will consider the *lifetime* of a particle before decay. For definiteness, we will discuss the neutron. Suppose that at $t = 0$ we have n_0 free neutrons (assuming $n_0 \gg 1$). In the course of time, the neutrons will become ever fewer due to decay, that is, $n(t)$ will be a decreasing exponential function

$$n(t) = n_0 e^{-t/\tau} \tag{1}$$

(here $e = 2.718\ldots$ is the base of the natural logarithm). The curve of this function is shown in Fig. 106. The constant τ, which has the dimensions of time, is called the *lifetime* of the neutron. This is the time during which

* *Free* neutrons are not stable, but atomic neutrons are stable. In unstable nuclei, neutron may decay, the phenomenon being called the β-radioactivity of atomic nuclei.

the number of neutrons will decrease e-fold. For neutrons $\tau = 10^3$ s.

It is remarkable that the lifetime τ does not characterize the time of existence of an individual neutron. An individual neutron may live just a minute or a day. It is in principle impossible to foresee the moment at which a concrete neutron will disintegrate. It only makes sense to speak about the *probability of decay*. The factor $e^{-t/\tau}$ on the right-hand side of (1) is the factor describing the probability for individual neutrons to decay during time t. It is immaterial in that case how long the neutron has lived by $t = 0$; for all the neutrons, the probability that they may live for time t is absolutely the same. It can be said that neutrons do not age.

Despite the fact that an individual neutron may in principle live indefinitely, the number of neutrons in a large ensemble falls off with time following a definite law. Speaking about the lifetime, we do not mean the lifetime of an individual neutron, but the time during which the total number of neutrons decreases markedly (to be more exact, e-fold).

How then do we treat the very act of decay? One should not believe that if a neutron decays into a proton, an electron, and an antineutrino, it means that before the decay the neutron constituted some combination of the above particles. The decay of a subatomic particle is by no means a decay in the literal sense of the word. This is an *act of transformation of the initial particle into some set of new particles* in which the initial particle *is annihilated* and new particles *are produced*.

It is worth noting that having discovered the β-radioactivity of atomic nuclei (the emission of electrons by nuclei), scientists decided at first that electrons enter the composition of nuclei. It was not until some time later that they understood that the electrons of β-radiation *are born at the time of decay* of neutrons of β-radioactive nuclei.

Another argument against the literal interpretation of the term "decay of a particle" is the fact that many particles may decay in *different* ways. So in the overwhelming majority of cases (more than 99 per cent), a positive pion decays following the above-mentioned scheme:

$$\pi^+ \to \mu^+ + \nu_\mu.$$

In other cases, however, it decays differently:

$$\pi^+ \to e^+ + \nu_e.$$

About a half of sigma hyperons Σ^+ disintegrate by the scheme

$$\Sigma^+ \to p + \pi^0,$$

whereas the other half by

$$\Sigma^+ \to n + \pi^+.$$

For a specific Σ^+-hyperon, it is impossible to predict how it will disintegrate, let alone the time of the disintegration.

Let us return to the lifetime of particles. As stated above, for the neutron this time is 10^3 s. For comparison we will quote the lifetimes: for muons, about 10^{-6} s; charged mesons, about 10^{-8} s; hyperons, about

10^{-10} s. As compared with neutrons, all these particles are like butterflies that live for only one day. And still, in terms of the microworld, they must be brought under the heading of *long-lived* particles. There are particles that live far shorter–10^{-16} s (the neutral pion) and even 10^{-23} s (the so-called *resonances*).

Worthy of special attention are the five subatomic particles that are stable and do not decay at all. Three of them (the photon and two neutrinos) move with the velocity of light in any inertial frame of reference. The infinite lifetime of these particles is, in essence, a consequence of the theory of relativity (see Chapter 8). More surprisingly, two particles with nonzero rest masses – electron and proton – also have infinite lifetimes. *Their stability is based on symmetry*, which is here expressed by specific conservation laws (see Chapter 13). It is to be stressed that the stability of the electron and the proton is crucial for the existence of stable atoms, and hence for our entire world.

Interconversions of Particles

In the world of elementary particles a wide variety of *interconversions* occur, as a result of which some particles are destroyed and others are born. *Decays of unstable particles* and *annihilations* in collisions of particles with antiparticles are examples of such interconversions.

Interconversions may also occur when *particles collide with particles*. The following are some examples of processes occurring when two protons collide with each other:

$$p+p \to p+n+\pi^+,$$

$$p+p \to p+\Lambda^0+K^+,$$

$$p+p \to p+\Sigma^+ +K^0,$$

$$p+p \to n+\Lambda^0+K^+ +\pi^+,$$

$$p+p \to p+\Xi^0+K^0+K^+,$$

$$p+p \to p+p+p+\bar{p}.$$

It can easily be found that the total rest mass of the particles born in these processes is larger than the double rest mass of a proton by a factor of 1.07, 1.36, 1.4, 1.43, 1.73, and 2, respectively. This suggests that in the above processes the kinetic energy of the protons involved in the collision must be sufficiently high. This energy goes into the production of the difference of the total intrinsic energies between the born and annihilated particles*. This difference is Δmc^2, where Δm is the difference of the total rest masses of the particles after and before the process. By increasing the kinetic energy of the protons, it is possible to observe the processes in which the number of particles born increases. In principle you can

* The intrinsic energy of a particle is the energy associated with its rest mass, that is, the energy mc^2. For more details see Chapter 13.

conjure up a really fantastic picture: two protons collide with enormous energy to produce a galaxy!

Suppose we wish to "split" protons by bombarding them with photons gradually increasing the energy of the latter. Instead of the splitting of protons we would observe various interconversions, for example the following:

$$\gamma + p \rightarrow p + \pi^0,$$

$$\gamma + p \rightarrow n + \pi^+,$$

$$\gamma + p \rightarrow p + \pi^+ + \pi^-,$$

$$\gamma + p \rightarrow p + p + \bar{p}.$$

This example demonstrates that interconversions make futile any attempts to split some particles by bombarding them with others. In actual fact, we observe *not the splitting of the particles bombarded but the birth of new particles.* New particles are born at the expense of the energy of the colliding particles.

Interactions of particles are studied in a chamber in which charged particles leave a distinct track. Widely used are chambers filled with liquid hydrogen in superheated state. A charged particle passing through the chamber causes the hydrogen to boil leaving a clearly visible track of small bubbles. Such chambers came to be known as *bubble chambers.* If the chamber is placed in a fairly strong magnetic field, the tracks of particles will be curved, and particles with opposite signs will curve into opposite directions.

Shown in Fig. 107*a* is a fascinating photograph of tracks taken in 1959 in the new liquid-hydrogen bubble chamber (72 inches). The chamber was bombarded by an antiproton beam. The picture shows a rare event–an antiproton on colliding with a proton produces a lambda hyperon and an antilambda hyperon

$$\bar{p} + p \rightarrow \Lambda^0 + \overline{\Lambda^0}.$$

The deciphering of this photograph is given in Fig. 107*b*. The antiproton collided with the proton at point *A*. Two electrically neutral (therefore invisible in the picture) particles – Λ^0 and $\overline{\Lambda^0}$ were born. At point *B* the antilambda hyperon disintegrated to yield an antiproton and a positively charged pion:

$$\overline{\Lambda^0} \rightarrow \bar{p} + \pi^+.$$

Notice that the tracks of these particles ($\bar{p}$ and π^+) diverge, which is due to the fact that their charges have opposite signs (the chamber had been placed into a magnetic field perpendicular to the plane of the photograph). At point *C* the antiproton collided with the proton, the annihilation following the conventional scheme:

$$\bar{p} + p \rightarrow \pi^+ + \pi^+ + \pi^- + \pi^-.$$

The tracks of the pions and the antipions diverge in the photograph. At point D the lambda hyperon decayed:

$$\Lambda^0 \rightarrow p + \pi^-.$$

Note that the events recorded in the photograph include *four* generations of particles. The first generation is represented by the initial antiproton; the second one, by the hyperon Λ^0 and the antihyperon $\overline{\Lambda^0}$. The third generation particles are the products of disintegration of

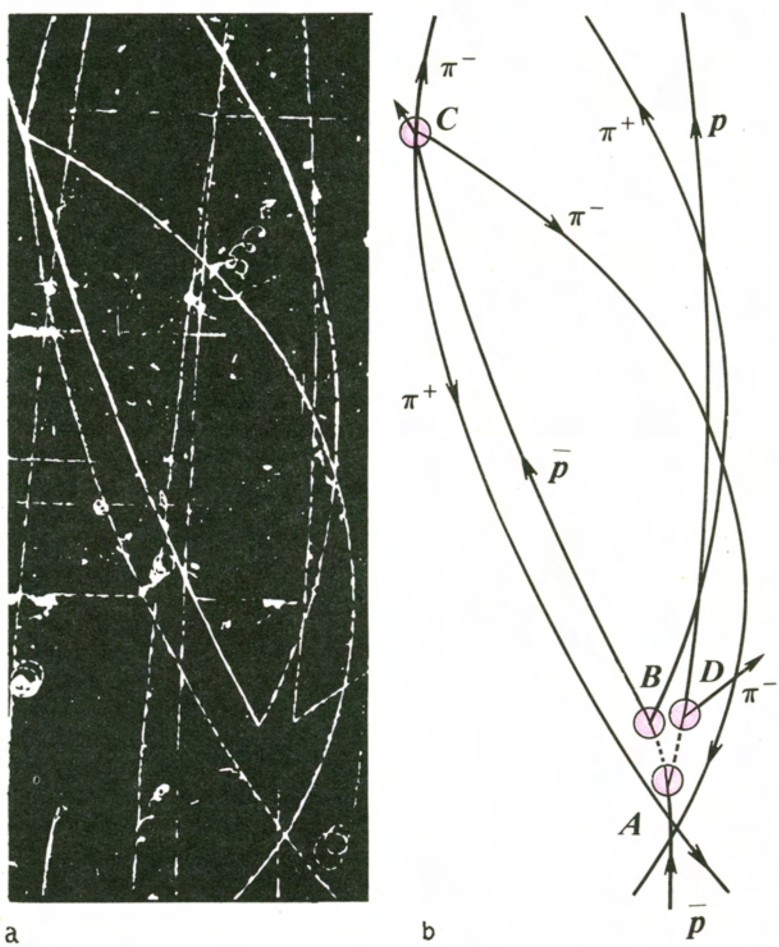


Fig. 107

Table 1 Elementary Particles

		Particle name	Particle symbol	Antiparticle symbol	Mass	Spin	Electric charge	Lifetime, s	Decay products
		Photon	γ	–	*0*	*1*	*0*	∞	–
LEPTONS		Electron	e^-	e^+	*1*	*1/2*	*−1*	∞	–
		Electron neutrino	ν_e	$\bar{\nu}_e$	*0*	*1/2*	*0*	∞	–
		Muon	μ^-	μ^+	*207*	*1/2*	*−1*	$2.2 \cdot 10^{-6}$	$(e^- \nu_\mu \bar{\nu}_e)$
		Muon neutrino	ν_μ	$\bar{\nu}_\mu$	*0*	*1/2*	*0*	∞	–
		Tauon	τ^-	τ^+	*3500*	*1/2*	*−1*	$\leq 10^{-12}$	$(e^- \bar{\nu}_e \nu_\tau)(\mu^- \bar{\nu}_\mu \nu_\tau)$ $(\pi^- \nu_\tau)$
HADRONS	MESONS	Pions	π^0	–	*264*	*0*	*0*	$0.8 \cdot 10^{-16}$	$(\gamma\ \gamma)$ $(\gamma e^- e^+)$
			π^+	π^-	*273*	*0*	*+1*	$2.6 \cdot 10^{-8}$	$(\mu^+ \nu_\mu)$
		Kaons	K^+	K^-	*966*	*0*	*+1*	$1.2 \cdot 10^{-8}$	$(e^+ \nu_e \pi_0)$ $(\mu^+ \nu_\mu)$ $(\pi^+ \pi^0)$
			K^0	$\bar{K}^0$	*974*	*0*	*0*	$K_s: 0.9 \cdot 10^{-10}$	$(\pi^+ \pi^-)(\pi^0 \pi^0)$
								$K_L: 5.4 \cdot 10^{-8}$	$(\pi^0 \pi^0 \pi^0)(\pi^0 \pi^+ \pi^-)$ $(\pi^- e^+ \nu_e)$
		Eta meson	η^0	–	*1074*	*0*	*0*	$\sim 10^{-17}$	$(\gamma\ \gamma)$ $(\pi^0 \pi^0 \pi^0)$ $(\pi^0 \gamma\ \gamma)$ $(\pi^+ \pi^- \pi^0)$
	BARYONS	Proton	p	$\bar{p}$	*1836.1*	*1/2*	*+1*	∞	–
		Neutron	n	$\bar{n}$	*1836.6*	*1/2*	*0*	*960*	$(pe^- \bar{\nu}_e)$
		Lambda hyperon	Λ^0	$\bar{\Lambda}^0$	*2183*	*1/2*	*0*	$2.5 \cdot 10^{-10}$	$(p\pi^-)$ $(n\pi^0)$
		Sigma hyperons	Σ^+	$\bar{\Sigma}^+$	*2328*	*1/2*	*+1*	$0.8 \cdot 10^{-10}$	$(p\pi^0)$ $(n\pi^+)$
			Σ^0	$\bar{\Sigma}^0$	*2334*	*1/2*	*0*	10^{-14}	$(\Lambda^0 \gamma)$
			Σ^-	$\bar{\Sigma}^-$	*2343*	*1/2*	*−1*	$1.5 \cdot 10^{-10}$	$(n\pi^-)$
		Xi hyperons	Ξ^0	$\bar{\Xi}^0$	*2573*	*1/2*	*0*	$3 \cdot 10^{-10}$	$(\Lambda^0 \pi^0)$
			Ξ^-	$\bar{\Xi}^-$	*2586*	*1/2*	*−1*	$1.7 \cdot 10^{-10}$	$(\Lambda^0 \pi^-)$
		Omega hyperon	Ω^-	$\bar{\Omega}^-$	*3273*	*3/2*	*−1*	$1.3 \cdot 10^{-10}$	$(\Xi^0 \pi^-)$ $(\Xi^- \pi^0)$ $(\Lambda^0 K^-)$

second-generation particles. Lastly, the fourth generation includes the particles born as a result of the annihilation of the proton and the secondary antiproton.

Interconversions of elementary particles enable us to gain a better understanding of the properties of the particles themselves. It is specifically these studies that made it possible to establish the existence of two types of neutrino (electron and muon neutrinos). Also, these studies made significant contributions to our understanding of conservation laws and invariance principles. All of these issues will be considered in later sections.

In conclusion, we will provide a table of elementary particles (Table 1). Note that the table contains two types of kaons K^0, which have significantly different lifetimes: *short-lived* kaons K_s and *long-lived* kaons K_l.

13 Conservation Laws and Particles

The strong hint emerging from recent studies of elementary particles is that the only inhibition imposed upon the chaotic flux of events in the world of the very small is that imposed by the conservation laws. Everything that can happen without violating a conservation law does *happen.*

K. Ford

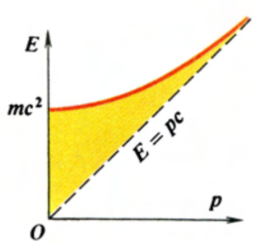


Fig. 108

Conservation of Energy and Momentum in Particle Reactions

In relativity theory the energy E of a particle having a rest mass m and momentum p is given by

$$E = \sqrt{(mc^2)^2 + (pc)^2}. \quad (1)$$

The dependence $E(p)$ is plotted in Fig. 108. At zero momentum ($p = 0$) the expression (1) yields the well-known *Einstein's formula*

$$E = mc^2.$$

The energy mc^2 is the *intrinsic* energy of a particle (the energy possessed by the particle in the frame associated with the particle). Suppose that the particle's momentum is not zero but small as compared with mc. In that case, the right-hand side of (1) becomes

$$mc^2 \sqrt{1 + \left(\frac{pc}{mc^2}\right)^2} \approx mc^2\left[1 + \frac{1}{2}\left(\frac{pc}{mc^2}\right)^2\right].$$

Hence

$$E \approx mc^2 + \frac{p^2}{2m}.$$

The first term describes the intrinsic energy of the particle, the second one is the well-known expression for kinetic energy. At sufficiently large momenta ($p \gg mc$), the expression (1) becomes

$$E \approx pc.$$

For particles with zero rest mass the relationship $E = pc$ clearly holds at any momentum (see the dash line in Fig. 108).

Suppose that a particle with rest mass m_1 decays into a particle with mass m_2 and a particle with mass m_3, the energy and momentum of these latter being E_2, $\vec{p}_2$ and E_3, $\vec{p}_3$, respectively. In the frame of reference associated with the initial particle the *laws of conservation of energy and momentum* have the forms:

for *energy*

$$m_1c^2 = \sqrt{(m_2c^2)^2 + (p_2c)^2} + \sqrt{(m_3c^2)^2 + (p_3c)^2};$$

for *momentum*

$$0 = \vec{p}_2 + \vec{p}_3.$$

It follows that for the decay to take place the following inequality must hold:

$$m_1 > (m_2 + m_3).$$

Or, in other words, *in a decay the total mass of the products must be smaller than the rest mass of the initial particle.*

Significantly, in a decay the magnitude of the momentum of a decaying particle is immaterial. No matter how we accelerate, say, a charged pion, all the same it can only decay into particles such that their total rest mass is smaller than the rest mass of the pion. With collisions the situation is different. By increasing the momentum of the particles involved in a collision, processes can be realized in which particles will be born with the total rest mass larger than that of the colliding particles.

Consider the process production of an electron-positron pair by collision of two photons

$$\gamma + \gamma \rightarrow e^- + e^+.$$

The initial particles here have no rest mass at all, nevertheless their collision produces two particles with rest mass.

Here is an example of *radiation-to-matter conversion.*

Let $\vec{p}_1$ and $\vec{p}_2$ be the momenta of the photons, and $\vec{p}_3$ and $\vec{p}_4$ the momenta of the electron and the positron. The laws of conservation will then read:

$$p_1 c + p_2 c = \sqrt{(mc^2)^2 + (p_3 c)^2} + \sqrt{(mc^2)^2 + (p_4 c)^2}$$

for *energy*, and

$$\vec{p}_1 + \vec{p}_2 = \vec{p}_3 + \vec{p}_4$$

for *momentum.*

It is easily seen that the total energy of the photons must be larger than $2mc^2$. If the photons with equal momenta collide head on, then $\vec{p}_3 + \vec{p}_4 = 0$. This means that an electron-positron pair may have essentially zero momenta, and so the required total energy of the photons appears to be minimal and equal to $2mc^2$.

The Conservation of Electric Charge and Stability of the Electron

In all the processes occurring in the world of elementary particles, the *law of conservation of electric charge holds: the total electric charge of the primary particles is exactly equal to the total electric charge of the secondary particles.*

The form of symmetry underlying this conservation law is more subtle

than those concerned with spatial and temporal translations, that is, the laws of conservation of energy, momentum, and angular momentum. But to formulate this principle we would have to resort to quantum mechanics. We will, therefore, confine ourselves to the remark that underlying the law of conservation of electric charge is the *symmetry of physical laws with respect to changes in the magnitude of the intensity of electric field.* The familiar statement that the magnitude of a potential has no physical meaning (and it is only the potential *difference* that matters), thus amounts to the statement that electric charge is conserved.

The stability of the electron is one of the most important consequences of the law of conservation of electric charge. Since the electron is a particle with *smallest* nonzero rest mass, its decay can only give rise to zero rest mass particles (recall that this conclusion follows from the laws of conservation of energy and momentum). But all the zero-mass particles are electrically neutral. Accordingly, the decay of an electron is forbidden by the law of conservation of electric charge.

The Three Conservation Laws and Neutrino

The existence of the neutrino had been predicted long before it was found experimentally, and the prediction was based on *energy conservation.*

It was found that the energy of an electron produced in the beta decay of a nucleus turns out to be different in different decay events and at all times it is smaller than half the total energy released in the process. Experimenters placed a beta-active sample within a heat-insulating lead chamber whose walls did not let in a single electron. Accuráte measurements have shown that the chamber heats up to a lesser degree than might be expected from the heat budget. It was suggested that in the beta fissure of nuclei the conservation of energy is invalid. The prominent Swiss physicist Wolfgang Pauli (1900-1958) came up with another explanation of the enigma of beta decay. Assuming that the *law of conservation of energy is also valid in the microworld*, in 1930 he came to the conclusion that in beta decay in addition to the electron some neutral particle is produced, which could not be recorded by the experimental set-up. It was this particle that carries away the energy that is obtained if from the energy liberated in the process we subtract the energy carried away by the electron. The famous Italian physicist Enriko Fermi (1901-1954), the author of the theory of beta decay, christened this neutral particle "neutrino", which is the Italian for "small neutron". So in the list of subatomic particles, appeared another entry, the neutrino. For a long time this particle had actually been a phantom, a particle only deduced from symmetry.

Neutrino was found experimentally in 1956. Frankly, by that time nobody questioned the very fact of the existence of the "elusive" neutrino. This was because the neutrino hypothesis was based not only on the law

of conservation of *energy* but also on the laws of conservation of *momentum* and *angular momentum*. Let us take a simple example, the beta decay of a neutron; the neutron decays following the equation mentioned above*

$$n \rightarrow p + e^- + \bar{\nu}_e.$$

We will consider this decay in the frame of reference associated with the neutron. For sufficiently slow neutrons, this frame is essentially the

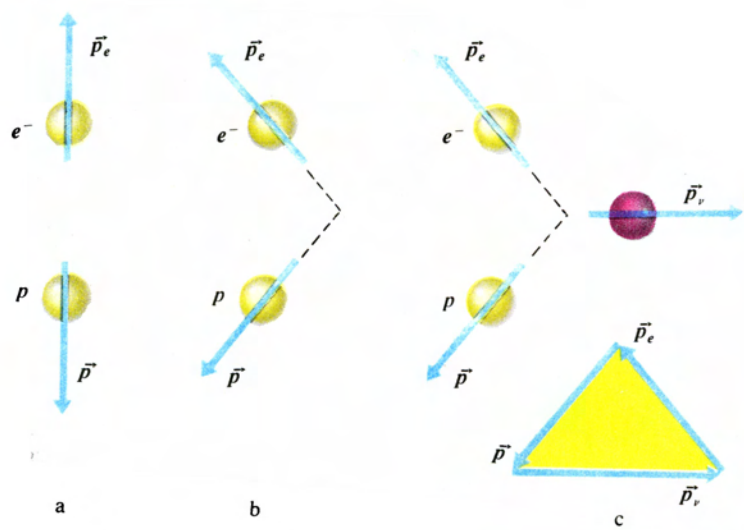


Fig. 109

lab frame. If this decay did not produce an antineutrino, then from *momentum conservation*, the proton and the electron would have to move off in opposite directions, as is shown in Fig. 109*a*. It was found, however, that the proton and the electron scatter in a different manner (Fig. 109*b*). This suggests that another particle is produced here whose momentum determined the observed picture (see Fig. 109*c*). The law of conservation of momentum dictates that the vector sum of the three momenta (the proton, electron, and third particle, that is, the antineutrino) is zero in the frame of reference associated with the neutron.

The *angular momentum* must also be conserved here. Before the decay the angular momentum was determined by the neutron's spin $s = 1/2$. If the decay products only consisted of a proton ($s = 1/2$) and an electron

* Note that the decay scheme includes not a neutrino but an antineutrino (electron). This was revealed later.

($s = 1/2$), the law of conservation of angular momentum would not hold in this case. In fact, the proton and the electron may have either parallel or antiparallel spins, and so the total spin may be either 1 or 0, and by no means 1/2. For the angular momentum to be conserved, another particle is necessary, such that its spin is 1/2. This explains Fig. 110, where the momentum vectors are shown in blue and the angular momentum vectors are shown in red (notice that the antineutrino is a right-handed helix).

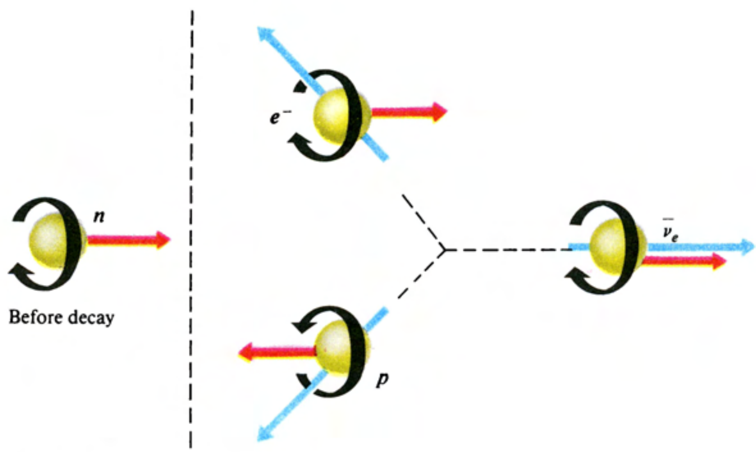


Fig. 110

Experimental Determination of Electron Antineutrino

In 1956 the American physicists Cohen and Reines obtained direct experimental proof of the existence of the antineutrino (and hence neutrino). They used the process

$$\bar{\nu}_e + p \rightarrow n + e^+.$$

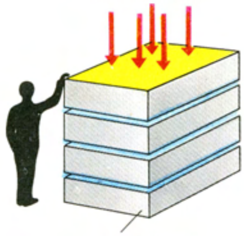
This sort of a process is highly unlikely. It is only known that the neutrino and the antineutrino interact with matter extremely weakly – they, essentially unhindered, pierce all barriers, the entire globe and even the Sun. To record such a particle it is necessary to use a sufficiently dense beam and advanced experimental techniques.

Cohen and Reines used an antineutrino beam from a high-capacity nuclear reactor. Into this beam they placed a special-purpose detector consisting of several layers of water separated by a scintillator capable of detecting individual photons (Fig. 111*a*). There is a tiny probability for an antineutrino to interact with a proton in the water to cause the above process that yields a neutron and a positron. It should be stressed that these events are, if any, exceedingly rare. It is for this reason, that the

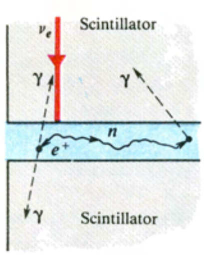
scintillator was utilized. A produced positron is stopped and undergoes an annihilation with an atomic electron

$$e^+ + e^- \rightarrow \gamma + \gamma.$$

As a result, two photons are created that fly off in opposite directions so that they can be recorded simultaneously in two adjacent scintillator layers (Fig. 111*b*). As for the neutron, it diffuses in a layer of water for a relatively long time (about 10^{-6} s), until it is caught by a cadmium nucleus (some cadmium is added to the water). After having absorbed the neutron, the cadmium nucleus emits a photon or photons, which are caught by a scintillator layer (Fig. 111*b*). Consequently, the scintillator must respond to an antineutrino-proton collision with three pulses: first a pair of simultaneous pulses are recorded using adjacent scintillators, and then, in about 10^{-6} s, another pulse. Both of the first two pulses correspond to a 0.5-MeV photon, and the third pulse to a 10-MeV photon. In Cohen and Reines's experiment this specific picture of pulses was actually observed (approximately three times per hour). Thereby the existence of the antineutrino was proved.



a



b

Fig. 111

At the time of the experiment just described, nobody suspected that there are two forms of neutrino (antineutrino), the electron and the muon neutrino. It is only natural that nobody knew that it was the electron neutrino that was found. The second type of the neutrino (muon) was discovered in 1962 by a group of experimentalists at Columbia University, USA. After it had been established that there exist two types of neutrino, two independent specific conservation laws were formulated – the *conservation of electron* and *muon numbers*.

Electron and Muon Numbers. Electron and Muon Neutrinos

The *electron* and *muon numbers* are specific charges of elementary particles. They have nothing to do with the electric charge. The *electron number* of the electron and the electron neutrino is assumed to be 1, and that of their antiparticles, -1. With all the other particles and antiparticles the electron number is zero. The *muon number* of the muon and the muon neutrino is 1, that of their antiparticles, -1. With all other particles (antiparticles) the muon number is zero.

In each process the total electron number of the reactants must be conserved. It follows, for example, that the creation of an electron must be accompanied with the creation of either an electron antineutrino or a positron, or with the annihilation of an electron neutrino:

$$\bar{n} \rightarrow p + e^- + \bar{\nu}_e, \quad \gamma + \gamma \rightarrow e^- + e^+, \quad n + \nu_e \rightarrow p + e^-.$$

Apart from the electron number, *in each process the total muon number must also be conserved.* So the creation of a muon must be accompanied by the creation of a muon antineutrino or the annihilation of a muon neutrino:

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu, \quad n + \nu_\mu \rightarrow p + \mu^-.$$

The *conservation of electron and muon numbers* in the decay of a muon dictates that in addition to an electron, an electron antineutrino and a muon neutrino must be produced:

$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu .$$

Specifically, the discovery of conservation of electron and muon numbers is associated with the solution of one problem that for a long time was unanswered – the so-called μ-e-γ-problem. It was noted long ago that nobody ever observed the muon reaction $\mu^- \to e^- + \gamma$. The answer came from the laws of conservation of electron and muon numbers: in this scheme *neither the electron nor the muon number is conserved*. This reaction thus appears to be *forbidden twice* (by two conservation laws).

The existence of *two independent conservation laws* (for electron and muon numbers) is closely related to the existence of *two different neutrinos* (*antineutrinos*). The electron neutrino (antineutrino) takes part in processes where an electron or positron is created or annihilated, whereas a muon neutrino (antineutrino) takes part in other processes, ones in which a muon or an antimuon is created or annihilated.*

That the electron and the muon neutrinos are two absolutely different particles was established in 1962 in the aforementioned experiment at Columbia University. A powerful beam of high-energy protons was directed from an accelerator at a target to produce a beam of pions π^+ and antipions π^-. Pions and antipions are known to decay in 99 per cent of all cases into antimuons μ^+ and muons μ^-, and hence muon neutrinos and antineutrinos:

$$\pi^+ \to \mu^+ + \nu_\mu, \quad \pi^- \to \mu^- + \bar{\nu}_\mu .$$

With the help of a 10-m thick iron wall, all the particles were trapped except for the above-mentioned muon neutrinos and antineutrinos. Into a flux of these neutrinos and antineutrinos a hydrogen-containing target was placed to study the products of processes occurring in rare collisions of a muon antineutrino with a proton. If the muon antineutrino were identical with the electron one, the following processes might be observed with equal probability:

$$\bar{\nu}_\mu + p \to n + \mu^+ \quad \text{and} \quad \bar{\nu}_e + p \to n + e^+ .$$

This was not the case, however. During 300 hours the experimenters recorded 30 tracks of antimuons μ^+ and found no sign of positrons. Thereby the existence of two different types of neutrino and antineutrino was clearly established.

---

* If the existence of the third type of neutrino (tauon) will be established experimentally, we will have to introduce another conserved number, the tauon number.

The Baryon Number and Stability of the Proton

We will now turn to the various processes involving baryons but not antibaryons. It was found that in these processes the number of baryons is always unchanged. For example, in the process

$$n \rightarrow p + e^- + \bar{\nu}_e$$

the baryon n decays just to give birth to another baryon p. In the process

$$p + p \rightarrow n + \Lambda^0 + K^+ + \pi^+$$

two baryons p annihilate and two baryons (n and Λ^0) are produced. Consequently, the annihilation of some baryons is compensated for by the production of others, the total number of baryons remaining the same.

Let us assign a specific number to each baryon, which is assumed to be unity. We will call it the *baryon number*. Photons, both neutrinos, an electron, muon and mesons have no such number (or it is zero). The fact that the number of baryons in various processes remains unchanged can be viewed clearly as the *law of conservation of baryon number*.

Further, we will remember that there are antibaryons. If the baryon number of the baryons is unity, then for antibaryons it must be set at -1 (recall that the antiparticles have the opposite signs of all the numbers). From baryon number conservation, processes must occur with *pair* creation or annihilation of an antibaryon and baryon. Such processes are actually observed. For example,

$$p + \bar{p} \rightarrow \pi^+ + \pi^- + \pi^0, \quad p + \bar{p} \rightarrow \Lambda^0 + \bar{\Lambda}^0, \quad p + p \rightarrow p + p + p + \bar{p}.$$

The *law of conservation of baryon number* is at present considered to be well established. According to that law, *in any process the difference between the number of baryons and the number of antibaryons remains unchanged.* Accordingly, for the entire Universe too, the difference between the total number of baryons and the total number of antibaryons is unchanged. It is worth noting here that by *the law of conservation of electron number the difference between the number of particles of an electron family and the number of their antiparticles remains unchanged.* In consequence, *the law of conservation of muon number leaves unchanged the difference between the number of particles of the muon family and the number of their antiparticles.* Curiously, there is no similar law for photons or mesons.

If the stability of the electron stems from the law of conservation of electric charge, then the *stability of the proton follows from the law of conservation of baryon number*. Among the baryons, the proton has the smallest mass, therefore there cannot be any baryons among its decay product. If this were not the case, the decay of a proton would lead to an uncompensated annihilation of a baryon. But such a process is prohibited by the law of conservation of baryon number.

In the rather large family of baryons only the proton is a stable particle. All the remaining baryons (neutron and hyperons) are unstable; each of them disintegrates to produce a lighter baryon. As regards the

proton, according to the law of conservation of baryon number, it simply cannot decay into anything. The mesons, muon and electron, which lie along the mass scale, have a zero baryon number.

The world around us (and hence we as well) could not exist if protons and electrons were not stable. This goes to prove the exclusive role of conservation laws, specifically the law of conservation of baryon number and electric charge.

Discrete Symmetries. *CPT*-Invariance

We have considered seven conservation laws: for energy, momentum, angular momentum, electric charge and three numbers (electron, muon, and baryon). The first four are related to the well-known properties of the symmetry of physical laws. One can expect that the remaining conservation laws (for electron, muon, and baryon numbers) express some symmetries, but we do not yet know which symmetries.

The above conservation laws are said to be *absolute* laws: they hold true at all times, in all the transformations of elementary particles*. To the seven absolute conservation laws we must add the eighth one – the *law of conservation of CPT-symmetry.*

CPT-symmetry is a conservation law for the combination of three sufficiently clear symmetries: the *symmetry with respect to the replacement of all particles by their respective antiparticles* (the so-called *charge-conjugation symmetry* or *C*-invariance), *mirror symmetry* (generally called *P*-invariance), *symmetry in time*, that is, *symmetry with respect to time reversal* (the so-called *T*-invariance).

CPT-symmetry means that if all particles were *simultaneously* replaced by appropriate antiparticles and mirror-reversed, and then the sense of time were reversed, the laws of physics would remain unchanged and all the physical processes would proceed as before. This statement is usually referred to as the *CPT-theorem.* The *CPT*-theorem is so firmly entrenched in the foundations of physics that, if it were to turn out not to be true, physical theory would be in shambles. "All hell will break loose", was how Abraham Pais once expressed it.

Turning to the symmetries underlying this conservation law, we note first of all that the symmetries we dealt with earlier in the book are associated with the conservation of energy, momentum, angular momentum, and electric charge, and they are *continuous* symmetries. In all of them a change that leaves the laws of physics unchanged can be made *arbitrarily* small; this change can be brought about smoothly, gradually, in other words, continuously. Apart from continuous symmetries there exist other symmetries that are considered in relation to changes that inherently cannot be continuous. They are associated with jumps, or are said to be *discrete.* All three symmetries involved in the *CPT*-theorem are *discrete.* It is clear that an object cannot be partially reflected in a mirror; the reflection is either possible or not. Likewise, one cannot

---

* Later in the book we will get acquainted with the laws that are not absolute. They hold in some transformations and do not in others.

replace a proton by an antiproton partially, the replacement either occurs or does not. The same is true of the reversal of the sense of time.

C-symmetry was covered in Chapter 12, *P-symmetry* in Chapter 9. We only have to discuss *T-symmetry*.

Everyday experiment indicates that time flows in one direction only. It might appear that in the world around us there is no invariance under time reversal. If we run a motion-picture film backward, events on it will look grotesque: people will walk backwards, the fragments of a broken vase will come together to form a whole vase, a swimmer will not dive into the water but, on the contrary, he will be expelled from the water feet first, and so forth.

An examination of transformations of subatomic particles shows, however, that *both senses of time are physically equivalent*. So, besides the process $e^- + e^+ \to \gamma + \gamma$ the reversed process is possible: $\gamma + \gamma \to e^+ + e^-$. And apart from the process $p + p \to p + n + \pi^+$ the reversed process $p + n + \pi^+ \to p + p$ is possible. True, "possible" by no means implies "equiprobable". In the last example, the reversed process is less likely than the direct one. This stems from the small probability of the meeting of three particles at once.

It is the low probability of reversed processes that accounts for the apparent noninvariance of physical laws under mental time reversal. The laws of physics as such are symmetrical with respect to the future and the past. But for any specific chain of events, a certain sequence, as a rule, turns out to be more *likely* than the opposite order.

Returning to *CPT*-symmetry, we will note that this *means CPT-invariance*, that is, invariance under three simultaneous operations: replacement of particles by antiparticles, mirror reflection and time reversal. In this connection, recall that energy can be defined as a quantity whose conservation is a consequence of invariance under a shift in time, momentum is a quantity conserved as a consequence of invariance under spatial translations, angular momentum is a quantity conserved as a consequence of invariance under spatial rotation (see Chapter 11). *CPT-symmetry* is the *product of three quantities: charge-conjugation parity* (*C-parity*), *space parity* (*P-parity*)*, and *time parity* (*T-parity*). Each of these is a conserved quantity that corresponds to some discrete symmetry, namely charge conjugation, mirror reflection, and time reversal.

It is only natural that the issue of *conservation laws for charge-conjugation parity, space parity, and time parity* presents itself. It has been found recently that, unlike *CPT*-symmetry, these conservation laws are not absolute. This interesting issue will be the subject of the following chapter.

---

* Space parity is often referred to as just parity.

14 The Ozma Problem

Assume we have already established fluent communication with Planet X. How can we communicate to Planet X our meaning of left and right? Although an old problem, it has not yet been qiven a name. I propose to call it the Ozma problem.

M. Gardner

What Is the Ozma Problem?

In 1900 the American writer of children's books, Lyman Baum, wrote his famous book *Wonderful Wizard of Oz*. The land of Oz was ruled by prince Ozma. Another of Baum's characters was a servant called the Long Eared Hearer who could hear sounds thousands of miles away. In 1960 the American astronomer Frank Drake started a project of using a powerful radio telescope to search for radio messages from the Galaxy in hope of picking up signals from intelligent inhabitants of distant planets. A long-time admirer of Baum and his Oz books, he called his project Ozma and his radio telescope the Long Eared Hearer. This story is behind the term the 'Ozma problem' suggested by Gardner.

Suppose that we exchange radio messages with inhabitants of some distant planet. Our signals are certain coded pulses, that is, sequences of pulses of various intensities. Using the universal laws of mathematical logic as well as the laws of physics which apply to the entire Universe, we can arm ourselves with patience and achieve a measure of understanding with extraterrestrials. If, for instance, we were to transmit a sequence of numbers representing the masses of nuclei of helium, lithium, beryllium, boron, carbon, etc., divided by the proton mass, we could expect that the extraterrestrials would guess that this sequence describes the periodic system of elements. After all, the ratios of nuclear masses to proton mass are the same throughout the Universe.

It is rather tempting to convey to the inhabitants of other planets visual images in the form of plane (two-dimensional) figures. Suppose we send out a sequence of pulses that is a coded description of the simple figure shown in Fig. 112*a* – a rectangular figure open in its right side. To begin with, we ask our extraterrestrial correspondent to prepare a rectangle divided into twenty square units – five lines with four units per line. By scanning the figure from top to bottom, left to right (in accordance with the numbering of units in Fig. 112*a*), we send out the sequence of pulses shown in Fig. 112*b*: a more intensive pulse corresponds to a darker unit. We ask the distant correspondent to copy our operations on his rectangle: to scan the units line-by-line left-to-right and colour them according to the sequence of pulses transmitted.

And here emerges a fundamental problem: our correspondent has no idea as to what we understand by *left*- and *right-handedness*, and so he does not understand what is meant by "to scan a line from left to right". If

he hits upon the right direction of scanning, he will clearly end up with a contour with a gap in the right side. If he scans the lines in the opposite direction, he will end up with a contour with a gap in the left side, not in the right one (Fig. 112*c*). It is unknown in what direction the extraterrestrial will actually scan his rectangle, and so it is unknown which (*left* or *right*) figure he obtained.

To be sure, to explain our concept of handedness it would be more convenient to transmit some kind of object that possesses reflection

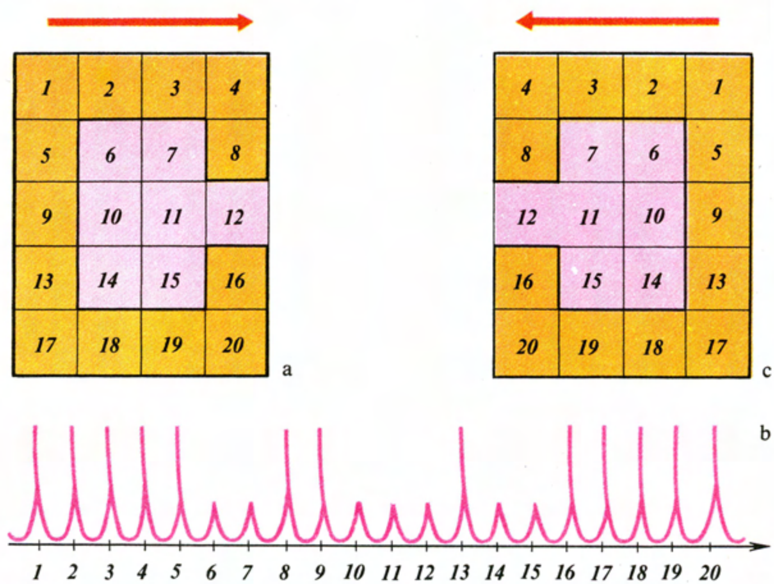


Fig. 112

asymmetry, for instance, a right screw. This is, however, absolutely impossible, since we can only make use of radio communication. He might be asked to look at some asymmetric constellation in the skies. But constellations do not look the same when observed from Earth and when observed from some distant planet.

There is thus no asymmetric object, no asymmetric structure which we could examine together with our correspondent from space. And so the question presents itself: *under these conditions, is there any way of getting across to the planet our concept of left and right*? It is this that is known as the Ozma problem. It is one of the most challenging and exciting problems of *communication theory*. This was conceived long before humanity began speculating about contacts with extraterrestrial civilizations.

The Ozma Problem Before 1956

If we had to explain the meaning of left and right to an Earth dweller, not an extraterrestrial, we would only have to say that rotation from left to right corresponds to the motion of the hands of a clock. But for an extraterrestrial this explanation will not do. There is no knowing the direction in which the hands (if any) of an extraterrestrial clock rotate.

Many natural compounds are known that turn the polarization plane of a light beam passed through them always in the same direction: to the right or to the left. This is because some compounds occur naturally on Earth only in the form of *certain* (left or right) *stereoisomers*. One can expect, however, that under extraterrestrial conditions these compounds occur in the form of other stereoisomers than those of Earth.

The animate world abounds in helices of *either handedness* (see Chapter 7). A wide variety of biological spirals are, however, of no help here. After all, the fact that all living things on Earth have their DNA molecules twisted only to the right, by no means suggests that in extraterrestrial beings they form right helices also.

The many manifestations of mirror asymmetry are not sufficient to solve the Ozma problem. *It is necessary for the vertical asymmetry to manifest itself in the very laws of nature.* Physicists have long thought that all the natural laws, without exception, are invariant under mirror reflection. This amounts to admitting that the Ozma problem is insolvable in principle.

Handling one problem in particle physics, the American physicists Lee and Yang in 1956 put forward the hypothesis that *space parity fails in processes of particle decay*. Chien-Shiung Wu staged an experiment to test the Lee-Yang hypothesis. The upheaval came on 15 January 1957 when it was reported that in particle decays the laws of nature are not invariant under reflection.

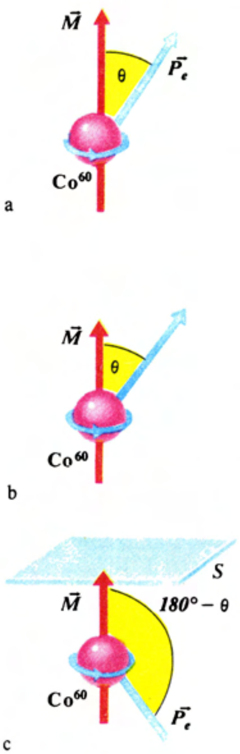


Fig. 113

The Mirror Asymmetry of Beta-Decay Processes

The Wu experiment studied the beta-decay of radioactive Co^{60}: $Co^{60} \rightarrow Ni^{60} + e^- + \bar{\nu}_e$. Let $\vec{M}$ be the intrinsic (spin) angular momentum of the cobalt nucleus, $\vec{p}_e$ is the momentum of the electron born in the decay, ϑ is the angle between $\vec{M}$ and $\vec{p}_e$ (Fig. 113*a*). Figure 113*b* depicts the reflection of the process of Fig. 113*a* in mirror *S*. Recall that the angular momentum $\vec{M}$ is an *axial* vector, and the momentum $\vec{p}_e$ is a *polar* vector; therefore, $\vec{M}$ and $\vec{p}_e$ are reflected in a *different way* (see Chapter 10 and Fig. 96). As a result, the angle between the vectors changes. If before the reflection it was ϑ, then after the reflection it will be $180° - \vartheta$ (Fig. 113*b*). The idea of the Wu experiment is fairly simple. In the decay of a Co^{60} nucleus, an electron may fly off at any angle ϑ to the direction of the nuclear angular momentum. If the beta-decay process is invariant under reflection, then the probability for the electron to be shot out at angle ϑ must equal that for $180° - \vartheta$, since the processes in which an electron is emitted at ϑ and at $180° - \vartheta$ are *mirror reflections of each other*. If these probabilities turn out to be different, then we have a violation of the

invariance of the process under mirror reflection. Accordingly, it is necessary to measure and compare the above-mentioned probabilities. To measure the probability for the electron to be shot out at one angle or another, it is necessary to consider the decay of a sufficiently large number of Co^{60} nuclei. Under normal conditions spin momenta of nuclei are oriented chaotically; in this case it was only required that for the majority of nuclei in a sample these momenta be oriented in a certain direction. It was this that constituted the main practical difficulty. To this

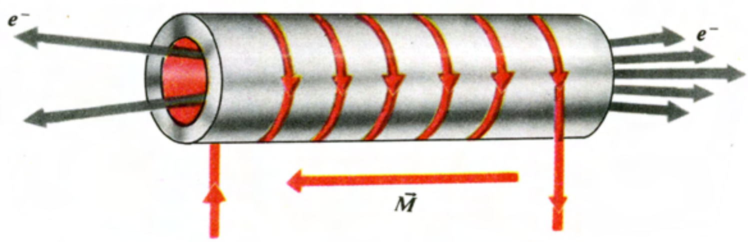


Fig. 114

end, Madam Wu cooled the cobalt sample nearly to the absolute zero of temperature ($T < 0.03$ K) to reduce all the joggling of its molecules caused by heat and placed it in a strong magnetic field. As a result, most nuclei in the sample aligned with the applied field. It then remained only to measure and compare the numbers of electrons emitted *along* and *opposite the field* (N_1 and N_2, respectively). If there was mirror symmetry, these numbers had to be equal. However, the experiment indicated convincingly that $N_2 > N_1$. It turned out that the *probability of an electron to be shot out against the nuclear spin is larger than along this direction.* Thereby it was proved that the beta-decay of Co^{60} nuclei has no mirror asymmetry. Later, other experiments were carried out to study other beta-decay processes. In all of them the invariance of natural laws under mirror reflection (or *P*-invariance) was found to be violated.

The Mirror Asymmetry in Decay Processes and the Ozma Problem

Now, it seemed the Ozma problem was solved. So, to explain to an extraterrestrial the meaning of left and right, we can do the following. We may ask him to manufacture a solenoid, place into it a cooled sample of radioactive cobalt and to count the electrons emerging from either solenoid end. We will then ask our space correspondent to align the solenoid so that his eye looks along its axis in the direction of the maximal emission of electrons (Fig. 114). In this case, the direction of the motion of electrons along the coils will be from left to right (as the hands of a terrestrial clock).

Note that here the direction of motion of conditional positive charges in the conductor will correspond to the direction from right to left, and

according to the right-handed screw, the ordering magnetic field (and hence the vector of the spin moment of cobalt nuclei) will be directed to the observer.

The Fall of Charge-Conjugation Symmetry

CPT-invariance (see Chapter 13) suggests that if any of the three invariances (here *P*-invariance) is violated, then at least one more invariance must be violated as well. It turned out that *apart from mirror symmetry, charge-conjugation symmetry* (*C*-invariance) *must also be violated.* In other words, nature's laws show noninvariance not only under replacement of left to right, but also under replacement of particles by antiparticles.

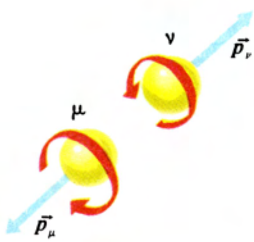


Fig. 115

We will illustrate this by turning to the case of muon neutrino and muon antineutrino. As noted (Chapter 12), the neutrino is like a left-handed helix, and the antineutrino, like a right-handed one. It is easily seen that this model of neutrino presupposes a violation both of *P*-invariance and *C*-invariance.

We will write the decay process for a pion π^+ in the form that takes into account the left-handedness of neutrino:

$$\pi^+ \to \mu_L^+ + \nu_{\mu L}. \qquad (1)$$

Since in the decay of a pion the antimuon and the neutrino are shot out in opposite directions (in the rest frame for the pion), then because of angular momentum conservation, both particles must appear as helices of the same handedness – in this case as left-handed helices (Fig. 115). If the process possessed a mirror symmetry, then apart from the decay according to (1), since in mirror reflection left-handed helices turn into right-handed ones, we would also have the decay process

$$\pi^+ \to \mu_R^+ + \nu_{\mu R}.$$

But such a decay is impossible, since the neutrino may only be a *left-handed* helix. If the charge-conjugation symmetry were not violated, then in addition to (1) we would have

$$\pi^- \to \mu_L^- + \bar{\nu}_{\mu L},$$

which is also impossible, since the antineutrino is a *right-handed* helix. Curiously enough, if we carry out a mirror reflection and a particle-to-antiparticle replacement *simultaneously*, then instead of (1) we will have the decays

$$\pi^- \to \mu_R^- + \bar{\nu}_{\mu R},$$

which are actually observed. This example may be used as an illustration of the interesting idea put forward in 1957 by the Soviet physicist L. Landau (1908-1968) and independently by Lee and Yang. They suggested that the so-called *combined parity* (*CP-parity* or a product of *C*- and *P*-parities) is conserved.

Combined Parity

Let us return to the beta-decay of cobalt-60 nuclei in an external magnetic field. It was shown in Fig. 113 how the process is affected by mirror reflection. In addition to mirror reflection, Fig. 116 also takes into account the replacement of the particles by their antiparticles. This figure includes four positions. A is the initial position in which the spin moment of the cobalt nucleus ($\vec{M}$) is aligned with the external magnetic field ($\vec{H}$ is the magnetic field intensity). B is the reflection of A in plane S;

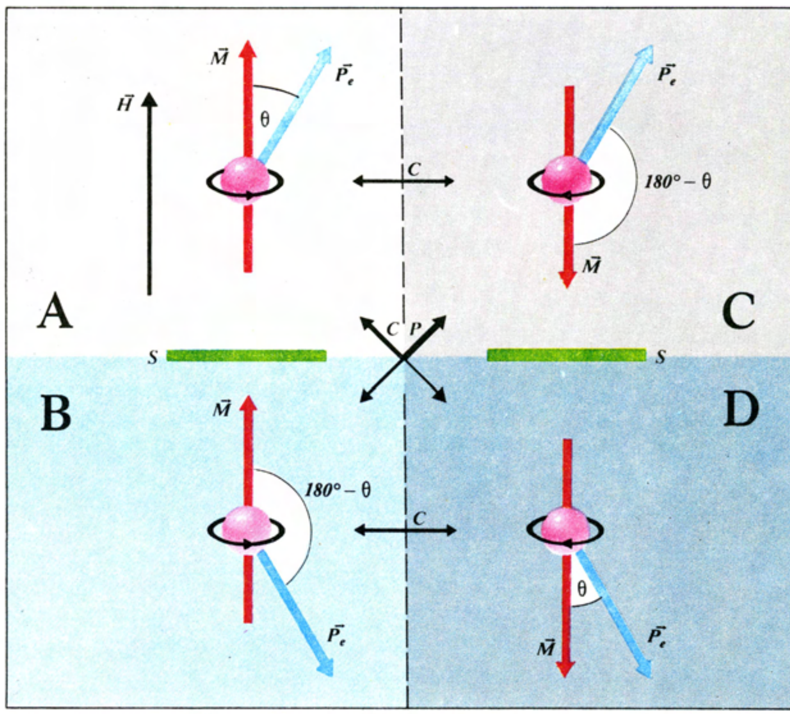


Fig. 116

this leaves the direction of $\vec{M}$, just like that of $\vec{H}$ (both are *axial* vectors), unchanged, whereas the direction of $\vec{p}_e$ (*polar* vector) changes. As a result the angle ϑ between $\vec{M}$ and $\vec{p}_e$ in the initial position becomes $180° - \vartheta$ in B. We have already discussed this.

Now suppose that we replace the particles by their respective antiparticles: the cobalt nuclei are replaced by cobalt antinuclei, the electrons are replaced by positrons. Unlike a nucleus, an antinucleus has a negative electric charge, and so an external magnetic field will align the spin moments $\vec{M}$ in the opposite direction to $\vec{H}$ (see C in Fig. 116). Reflecting in S gives D.

It is easily seen that if we carry out *simultaneously* the two transformations – *mirror reflection* and *particle-to-antiparticle replacement* – we obtain either the transition $A \equiv D$ or $B \equiv C$. This leaves the angle between $\vec{M}$ and $\vec{p}_e$ unchanged, or leaves invariant the decay process in question. This invariance is known as *combined* or *CP-invariance*. Accordingly, they attest to the *conservation of combined parity* (*CP-parity*).

Conservation of combined parity implies that the laws of nature remain unchanged not when we venture into the looking-glass land, and not when we venture into the antiworld, but when we cross over into the *looking-glass antiworld*. Put another way, the laws of physics turn out to be symmetrical under reflection in that imagined mirror, which at the same time effects a replacement of particles by their respective antiparticles (and antiparticles by particles). In the words of the Soviet physicist Ya. Smorodinsky, "*some test will be passed with the same result both by a left-handed screw made from matter and a right-handed helix made from antimatter*".

From the conservation of combined parity, in decay experiments the *mirror asymmetry comes from the fact that in our corner of the Universe particles outnumber antiparticles markedly*. Our world can be said to be asymmetrical with respect to mirror reflection only because it is asymmetrical in terms of the density of particles and antiparticles. Imagine that in the world the density of particles and antiparticles were the same (one has to suppose at the same time that the processes of annihilation of particles and antiparticles are forbidden in some sort of way). If that had been the case, Wu would have obtained a symmetrical result. In fact, in an external magnetic field half the nuclei (neutrons and protons) in the cooled sample of cobalt (anticobalt) would tend to align with the field, and half the nuclei (antineutrons and antiprotons) against the field. In this case, we would have an equal mixture of positions A and C in Fig. 116. Clearly, reflecting yields the same mixture.

Combined Parity and the Ozma Problem

Since the laws of nature do not enable us to distinguish a *left-handed helix of matter* and *a right-handed helix of antimatter*, this suggests that the violation of spatial parity in decays of particles by no means solves the Ozma problem. It is of no help to attempt to use the Wu experiment to explain to an extraterrestrial our understanding of left and right, if we do not know a priori which he is made of, matter or antimatter. If he lives in an antiworld, then by repeating the Wu experiment and using our explanations, he would consider right what we consider left. The fact is that in an antiworld the solenoid coils carry positrons, not electrons, the electric charge of nuclei is negative, not positive.

We can imagine the following fantastic situation. We have agreed with our extraterrestrial to meet in space and set out to our rendezvous. And so, having put on our space-suits, we walk out of our ships and move toward each other. You extend your right hand and suddenly you see that he, aware of the terrestrial custom of handshaking, extends his left

hand, not his right one. Do not touch it because you are facing a dweller of an antiworld.

To sum up, the conservation of combined parity gives the Ozma problem a new perspective. *To solve this problem we will have to find out beforehand whether our companion is made of matter or antimatter.*

The Solution to the Ozma Problem

To understand with which world (a conventional or antiworld) we are in contact, apart from radio signals we can make use of the neutrino communication channel. Our Sun is a source of neutrino; that is, of left-handed helices. It is, therefore, in principle sufficient to send to our distant correspondent a solar neutrino and ask him to compare its handedness with that of the neutrino sent out by the luminary in the world of the extraterrestrial. Unfortunately, we at present cannot even think of a way of sending our neutrino to a distant planet.

And still there is a solution to the Ozma problem. Leaving out details, which can be far too complicated for this book, we note that physicists have revealed that in one decay *CP-parity is not conserved.* Considering *CPT*-invariance, this means that in this process *T-parity must also not be conserved* (that is, the symmetry with respect to time reversal must be violated). As it has been noted above, there are two types of *neutral kaons* – long-lived kaons (K_l) and short-lived kaons (K_s). The former live about 10^{-8} s, the latter, 10^{-10} s. In accordance with the conservation of combined parity, kaons K_l decay into three pions, and kaons K_s into two pions (the decay schemes are given in the table "Elementary particles"). In 1964 it was found that every now and then (in about one case out of 1000) a kaon K_l decays into two pions, not three. This suggested that *for kaons the law of conservation of CP-parity is not completely exact.* It is this additional violation of symmetry in the laws of physics that enables, in principle, the Ozma problem to be solved, or rather to find of which material the extraterrestrial correspondent consists, matter or antimatter.

We will have to ask him to observe the process of decay of neutral kaons, for example, to measure the density of the kaons K_l in the beam at various distances from the place of origin of the kaons and to report the results of his measurements. If the report comes from antiworld, then it will differ from the results of our measurements.

15 Fermions and Bosons

All the particles in nature are either fermions or bosons. Thus, there occur only antisymmetric or only symmetric states of the same particles.

P. Dirac

The Periodic Table and the Pauli Principle

The advances of atomic physics made it possible to provide a substantiation for the Periodic System. According to modern thinking, as the number of chemical elements increases, the electron shells are gradually filled up. The first to be filled up are shells that have the strongest binds with the atomic nucleus, that is, the closest to it. The first (the closest to the nucleus) shell takes only two electrons to be filled, whereas the second and third take eight electrons, and the fourth and fifth, eighteen, and so on. We thus obtain the sequence: 2, 8, 8, 18, 18, ... The number of the chemical elements in the first five periods of the Periodic Table are exactly like these.

The first shell may contain only two electrons, the second not more than eight. Why? The answer to this question is two-fold. First, to each shell there corresponds a *definite* number of possible states of the electron – two for the first shell, eight for the second shell, and so on. Second, *in each state there may be only one electron.* This means that in the atom you cannot find two electrons with the same characteristics that define its state, such as energy, orbital angular momentum, its projection, spin projection. And so at least one of these characteristics must be different.

This rule that forbids two or more electrons from occupying the same state is known as the *Pauli exclusion principle.* In the simplest terms, the principle can be understood as a rule that more than two electrons cannot reside in the place, and these electrons must have opposite spins (or must be in different spin states). Like any prohibition principle, the Pauli exclusion principle expresses a certain *symmetry of natural laws.* This is the so-called *commutative symmetry.*

Commutative Symmetry. Fermions and Bosons

This is the symmetry *with respect to commutation of any two particles of the same type*, specifically as applied to electrons. Physically, nothing changes if an electron in state *1* is placed in state *2*, and the electron in *2* is placed in *1*. This symmetry implies that all the electrons in the Universe are *identical.* Also identical are all the protons, all the neutrons, all the hydrogen atoms, all the oxygen atoms, and so on.

Our planet consists of about 10^{50} atoms. And this prodigious number of atoms consists of only several dozen varieties. Furthermore, we are confident that the entire Galaxy, the entire Metagalaxy, and the entire

Universe are constructed of several hundred various building blocks. All the chemical elements of the Universe can be arranged as a table that has about a hundred cells.

This, however, does not exhaust the profound meaning of commutative symmetry. *This symmetry dictates that all the particles in nature are split into two categories that behave differently in an ensemble.* One category obeys the rule that particles of the same type, for example electrons, must avoid one another. According to these rules, identical particles may only be alone in a state. All the particles of this category are known collectively as *fermions* (from the name of the Italian physicist Fermi). The other category is governed by precisely opposite rules, which not only allow but even dictate that identical particles concentrate densely in states. These particles are known as *bosons* (from the name of the Indian physicist Bose).

There is a connection between the spin s of a particle and its behaviour in an ensemble. All the particles with *half-integral* spin ($s = 1/2, 3/2, \ldots$) are fermions, whereas all the particles *without a spin* or *integer* spin are bosons. Apart from electrons, fermions include other leptons, and barions. All of them obey the exclusion principle formulated above for electrons: *if a state is occupied by a fermion, then no other fermion of this type may be in this state.* Other bosons are photons and mesons. In any given state there may be any number of the same bosons. Moreover, *the more densely is a given state populated the higher the probability that other bosons of this type will come to it.*

We have thus, on the one hand, clearly expressed individualism (leptons and barions), and, on the other hand, as clearly expressed collectivism (photons and mesons). In this connection it is worth noting that there is a marked difference between leptons and baryons, for one thing, and photons and mesons for the other. This is primarily explained by the fact that for the former there *exist* conservation laws according to which the difference between particles and antiparticles remains unchanged (conservation laws for electron, muon, and baryon numbers), whereas for the latter *there are no* such laws.

Taking the Periodic Table as an example, we can clearly see just how fundamental is the fact that electrons are fermions. It is exactly the fermion nature of electrons that accounts for the peculiarities of the population of the atomic levels by electrons. Should the Pauli exclusion principle suddenly fail to apply to electrons, then in that case and also in all atoms all the electrons would occupy the level with the lowest energy. And this would destroy the diversity of elements.

Note also that it is the fermion nature of electrons that does not allow atomic nuclei in a solid to approach one another too closely. At close distances the electron shells would overlap, in other words, many electrons would be in one place. But this is forbidden by the Pauli exclusion principle. As a result, the atoms remain separated by fairly decent distances from one another (about 10^{-10} m or more), which is no less than 10^3 times larger than the size of atomic nuclei.

Symmetrical and Antisymmetrical Wave Functions

It is to be stressed that a fermion does not admit other sister fermions to the state it occupies, just like a boson attracts other bosons. The fact *is in no way connected with any special forces* (repulsion or attraction) which act between the particles. The fermion or boson nature of particles is their fundamental property associated not with force interactions but with the *symmetry under commutations of particles.* Therefore, it would be instructive to explain even briefly how commutative symmetry leads to the presence of fermions and bosons in nature.

Note that in quantum mechanics the state of a microobject is described using some function called the *wave function.* It is significant that a physical meaning is attached not to the wave function but to its squared modulus, which describes the *probability of finding the microobject in a given state.*

Let $\psi_1(I)$ be the wave function of particle *I* in state *1*, and $\psi_2(II)$ is the wave function of particle *II* in state *2*. Consider the microobject as a system of particles *I* and *II*. The wave function of the microobject $\Psi(I, II)$ can be expressed as the product of the wave functions of the constituent particles. Since the particles are assumed to be identical, it is then unknown which of them is really in state *1* and which in state *2*. We will then have to take into account both $\psi_1(I)\psi_2(II)$ and $\psi_2(I)\psi_1(II)$ (as if the particles in the microobject were continually exchanging their places). *Commutative symmetry* requires that the wave function $\Psi(I, II)$ of the object meet the condition

$$|\Psi(I, II)|^2 = |\Psi(II, I)|^2.$$

Out of the combinations of the above products of single-particle wave functions we can construct two functions that meet this condition

$$\Psi_S(I, II) = \psi_1(I)\psi_2(II) + \psi_2(I)\psi_1(II)$$

and

$$\Psi_A(I, II) = \psi_1(I)\psi_2(II) - \psi_2(I)\psi_1(II).$$

The first of these is *symmetrical*, it *does not change* its sign under commutation: $\Psi_S(I, II) = \Psi_S(II, I)$. This function describes a system of *bosons.* The second function is *antisymmetric*, it *changes* its sign under commutation of particles: $\Psi_A(I, II) = -\Psi_A(II, I)$. This function describes a system of *fermions*. This can be readily verified. If we assume that both particles are in the same state, for example state *1*, then it follows from the expression for Ψ_A that this function vanishes. Hence this situation is impossible.

The Superfluidity of Liquid Helium. Superconductivity

At extremely low temperatures (under 2.19 K), He^4 forms a liquid that has a highly interesting property: its motion along a narrow capillary is characterized by the *total absence of viscosity.*

Liquid helium flows undergoing no resistance from the walls. This phenomenon is called *superfluidity.* This remarkable phenomenon comes from He^4 atoms being bosons. Note that a system consisting of an even number of fermions behaves as a boson. A commutation of two such systems amounts to a commutation of an even number of fermion pairs. A commutation of each pair of fermions changes the sign of the total wave function. If the sign changes an even number of times this means that it remains the same.

At very low temperatures, when the effect of the thermal motion of atoms, which scatters them over different states, becomes negligible, the rule manifests itself in full measure, which states that bosons must concentrate in one state. As a result, all the He^4 atoms concentrate in a state characterized by a definite momentum, and so they move along the capillary as a *single whole.* In that case, the liquid displays no viscosity: the presence of viscosity requires that different regions of the liquid travel with *different* velocities.

Unlike He^4, the atoms of He^3 are fermions. No wonder then, that when cooled down to 2 K, helium containing the isotope He^3 does not become superfluid. But the superfluid He^3 still exists. It was obtained in 1974 at stronger cooling – down to 0.0027 K. At such extremely low temperatures, a highly curious effect occurs in helium: He^3 atoms are *paired.* Each pair is clearly a boson. As a result we observe superfluidity.

Note one more curious phenomenon – the *superconductivity of metals.* It is well-known that at temperatures near absolute zero many metals begin to conduct electric current essentially without resistance. So, lead goes over to a superconductive state at 7.26 K, tin at 3.69 K, aluminium at 1.14 K, zinc at 0.79 K. The phenomenon of superconductivity can be viewed as the phenomenon of the superfluidity of the electron fluid formed in a metal by conduction electrons. The fact is that at low temperatures electrons *combine to form pairs* that behave like bosons. This is due to the interaction of electrons with crystal ions.

It is easily seen that superconductivity and superfluidity are essentially the same in nature. They are conditioned by the fact that He^4 atoms, pairs of He^3 atoms, and electron pairs are bosons.

Induced Light Generation and Lasers

In recent years great strides have been made in *quantum electronics*, a new area that came into life when an amazing light generator, the *laser*, was invented in 1960. The principle of the laser is the *induced emission* of light by matter.

The gist of this phenomenon is as follows. Suppose that in a substance we have excited atoms, each of which, when jumping back to the initial (unexcited) state can emit a photon with a definite energy. Under normal

conditions atoms accomplish such transitions in an *uncorrelated* manner, at different times, the photons emitted being sent out in *different* directions. This is called the *spontaneous* emission of light. One can, however, control the process to make the excited atoms return to the initial state at the *same time*, having emitted photons in a *definite* direction. This is the *induced* emission of light.

The phenomenon of induced emission is directly related to the boson nature of photons. A photon that flies past the excited atoms, when its energy is equal to that of the transition in the atoms, will *initiate* massive production of new photons in the *same* state in which it is itself. This phenomenon is used in the laser.

16 The Symmetry of Various Interactions

The state of the art in particle physics does not differ markedly from what you observe when you sit in a concert hall just before the performance begins. Many (but not all) musicians have appeared on the scene. They are tuning their instruments. At times you can hear some interesting musical passages: the improvisations are to be heard from all directions, sometimes wrong notes are also heard. The scene is pregnant with the expectation of the moment when the first sounds of the symphony will be heard.

A. Pais

The Principal Types of Interactions

According to modern views, there are *four* main types of forces in nature, or rather four types of interactions – *strong* (*nuclear*), *electromagnetic*, *weak*, and *gravitational*.

Nuclear forces strongly bind the neutrons and the protons in atomic nuclei. They are responsible for a wide variety of nuclear reactions, specifically for those reactions that release energy in the core of a nuclear reactor at an atomic power station. Hadrons are responsible for strong interactions (baryons and mesons), while leptons do not participate in them.

Electromagnetic interactions, it seems, are now the ones which occur most often: we encounter them when studying electric and magnetic phenomena and properties of matter and electromagnetic (in particular optical) radiation. These interactions determine the structure and properties of atoms and molecules. They encompass Coulomb forces, the forces acting on a current-bearing conductor, the forces of friction, resistance, elasticity, chemical forces, and what not. All the elementary particles, except for both neutrino and antineutrino participate in electromagnetic interactions.

Weak interactions are predominant in the realm of subatomic particles. They are responsible for the interactions of particles involving neutrino and antineutrino (specifically beta-decay processes). Furthermore, they are involved in neutrinoless decays that are characterized by a relatively long lifetime of the decaying particles – about 10^{-10} s or more*. These processes include the decays of kaons and hyperons.

Gravitational interactions are inherent in all particles, without exception, but they are of no significance for elementary particles. These interactions only manifest themselves on a sufficiently large scale when the masses involved are rather large. They are responsible, say, for the attraction of the planets to the Sun or for the falling of an object onto the ground. In what follows we will ignore gravitational interactions.

Interactions differ markedly in terms of the forces or energies involved. The strong interaction is about 100 times higher than the electromagnetic

---

* Speaking about the long lifetime of a particle, we compare it with the time during which light covers the distance of the order of the atomic nucleus itself, that is, $l \approx 10^{-15}$ m. The reference time taken to be a "unit time" in the world of elementary particles is about $l/c \approx 10^{-23}$ s.

one and 10^{14} higher than the weak one. *The stronger the interaction, the faster it carries out its task.* So the particles called resonances, whose decay occurs through nuclear interactions, have a lifetime of about 10^{-23} s; the neutral pions, which decay through an electromagnetic interaction ($\pi^0 \rightarrow \gamma + \gamma$), have a lifetime of 10^{-16} s; decays through a weak interaction have a lifetime of 10^{-8}-10^{-10} s. The strong interaction produces fast processes, the weak interaction, slow processes. The duration of a process is defined as a quantity that is the reciprocal of the probability of the process per unit time. The smaller the probability, the slower the process. It is to be recalled in this connection that the neutrino and antineutrino processes characteristic of the weak interaction are highly unlikely.

Unlike the electromagnetic interaction, strong and weak interactions manifest themselves over extremely short distances, or rather, have a *small range*. The strong interaction between two baryons or mesons only shows up when the particles approach each other and come within a distance of only 10^{-15} m. The range of the weak interaction is yet shorter, it is known to be within 10^{-19} m.

The most interesting difference between the types of interactions is associated with *symmetry*. All the interactions of particles are controlled by the absolute conservation laws discussed in Chapter 13. But there exist conservation laws (and pertinent symmetry principles), which are valid for some interactions and not for other interactions. So the laws of conservation of spatial and charge parity (P-invariance and C-invariance) hold for both the electromagnetic and strong interactions, but they do not hold for the weak interaction. There is the rule: *the stronger an interaction the more symmetrical it is.* Put another way, the weaker an interaction the less it is controlled by conservation laws. In the words of Ford, "*weaker interactions turn into infringers of the law, and the weaker an interaction the more lawlessness.*"

Isotopic Invariance of Strong Interactions. The Isotopic Spin (Isospin)

Suppose that all the protons in the atomic nucleus are replaced by neutrons, and all the neutrons by protons. The resultant nucleus is called the mirror nucleus of the initial nucleus. Mirror nuclei are, for example, the pairs: Be^7 and Li^7, B^9 and Be^9, C^{14} and O^{14}, and so on (Fig. 117). It was noticed long ago that pairs of mirror nuclei have similar properties: essentially the same nuclear binding energy, the similar structure of energy level, the same spin. The similarity of mirror nuclei reflects a measure of *symmetry of nuclear forces*, namely the fact that the nuclear forces between two protons are the same as the forces between two neutrons.

This symmetry is a special case of the so-called *isotopic invariance*. The latter means that from the point of view of the strong interaction, the system *p-p* (proton-proton) is identical with not only the system *n-n* (neutron-neutron) but also with the system *p-n* (proton-neutron). Stated

another way, *the nuclear forces are independent of the electric charge of particles.*

Associated with the isotopic invariance of the strong interaction is the concept of *isotopic spin* (*isospin*). It is worth recalling in this connection that the proton and the neutron can be viewed as two charge conditions of one particle – the nucleon (see Chapter 12). It is said that the proton and the neutron form an *isotopic doublet*. Isotopic doublets are also formed by two xi-hyperons (Ξ^-, Ξ^0) and two kaons (K^0, K^+). It

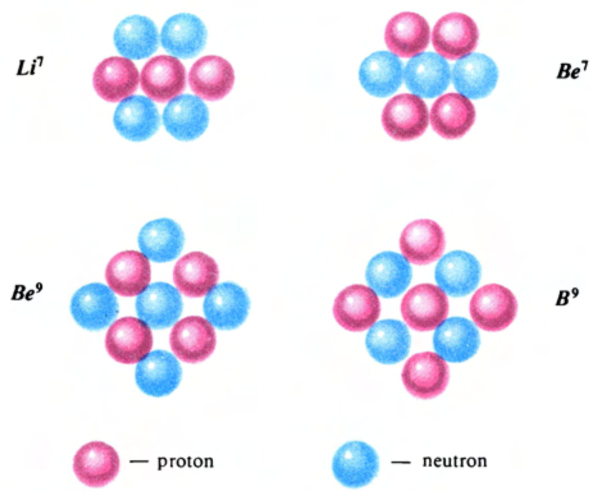


Fig. 117

appeared that pions are better combined into a *triplet*, by adding to π^+ and π^0 an antipion π^-. An isotopic *triplet* is also formed by three sigma-hyperons (Σ^-, Σ^0, Σ^+). As to the lambda-hyperon Λ^0, omega-hyperon Ω^- and eta-meson η^0, to each of them an isotopic *singlet* is put in correspondence. Isotopic multiplets of known elementary particles come in three types – triplets (three charge states), doublets (two charge states), and singlets (one charge state). True, there exists a multiplet with four charge states. This multiplet is formed by Δ-particles which belong to short-lived baryons, called resonances (Δ^-, Δ^0, Δ^+, Δ^{++}). The particle Δ^{++} has a positive electric charge, whose magnitude equals the doubled electron charge. Each isotopic multiplet is characterized by a quantity called the *isotopic spin* (*isospin*). The magnitude of isospin I of a particle is related to the number of charge states n in the multiplet by the relationship $n = 2I + 1$. Recall that the spin s of a particle is related to the number of spin states of the particle in exactly the same way. The analogy between the isospin and the spin, although formal (spin and isospin are physically absolutely different), is rather profound. One must recall that if the spin vector is in the conventional space, the isospin vector is

considered in some fictitious space (called *isotopic space* or *isospin space*). The electron's spin is $s = 1/2$, its projection in a given direction in conventional space takes on the values $s_z = +1/2$ and $s_z = -1/2$. The isospin of the nucleon (nucleon doublet) $I = 1/2$, its projection in some "direction" in the isospin space assumes the values $I_\zeta = +1/2$ (for the proton) and $I_\zeta = -1/2$ (for the neutron). Table 2 contains the various isotopic multiplets, as well as the values of the isospin I and the projections of the isospin I_ζ. Note that for an antiparticle the projection

Table 2

Isotopic Multiplets

Isotopic Multiplets	Singlets $n = 1$			Doublets $n = 2$			Triplets $n = 3$		Quadruple $n = 4$
	η^0	Λ^0	Ω^-	K^+ K^0	p n	Ξ^0 Ξ^-	π^+ π^0 π^-	Σ^+ Σ^0 Σ^-	Δ^{++} Δ^+ Δ^0 Δ^-
Electric charge	0	0	−1	+1 0	+1 0	0 −1	+1 0 −1	+1 0 −1	+2 +1 0 −1
Isospin projection	0	0	0	+1/2 −1/2	+1/2 −1/2	+1/2 −1/2	+1 0 −1	+1 0 −1	+3/2 +1/2 −1/2 −3/2
Isospin	0	0	0	1/2	1/2	1/2	1	1	3/2

of the isospin has the sign opposite to that of the corresponding particle. When dealing with the strong interaction of particles, the *isospin vectors of particles must be combined by the same rules as the spin vectors.* Specifically, the isospin projection of several particles is the algebraic sum of the isospin projections for individual particles.

The isotopic invariance of the strong interaction lies at the foundation of the physics of isospin formalism. This invariance means that the *laws of nature are invariant under rotations in isospin space.* This finds its expression in the *law of conservation of isospin* (just as the invariance of the laws of nature under rotations in conventional space is expressed in the law of conservation of angular momentum). In all the strong interactions of subatomic particles, the *total isospin of a system of particles is conserved.* Note that also conserved is the total projection of the isospin, which in fact implies that the total electric charge of the particles is conserved as well.

We will now illustrate the conservation of isospin using two processes: $p + p \rightarrow \pi^+ + D$ and $n + p \rightarrow \pi^0 + D$, where D is the deuteron (the nucleus of heavy hydrogen, which consists of a neutron and a proton). The deuteron's isospin is zero; therefore, the products of reactions in the general case have the total isospin that is equal to the isospin of pions, that is, unity. In the first reaction the sum of the isospin projections will be $1/2 + 1/2 = 1$, hence the isospin itself is 1. In the second reaction the total projection of the isospin is zero ($-1/2 + 1/2 = 0$); in this case the total isospin may be either unity or zero. Both values are equiprobable, therefore only in half of the cases a collision of a neutron with a proton can result in a reaction producing a deuteron. This suggests that the reaction $n + p \rightarrow \pi^0 + D$ must be half as likely as the reaction $p + p \rightarrow \pi^+ + D$. Experiment supports this prediction made on the basis of isospin conservation.

Strangeness Conservation in Strong and Electromagnetic Interactions

In the years 1947-1955 kaons and a number of hyperons (in collisions of pions with nucleons) were discovered. The particles discovered turned out to be fairly strange. First, they came in *pairs*–a kaon paired with a hyperon. For example,

$$\pi^- + p \rightarrow K^0 + \Lambda^0, \quad \pi^- + p \rightarrow K^+ + \Sigma^-, \quad \pi^+ + p \rightarrow K^+ + \Sigma^+.$$

Second, the lifetime of new particles produced without leptons ($K^+ \rightarrow \pi^+ + \pi^0$, $\Lambda^0 \rightarrow p + \pi^-$, $\Lambda^0 \rightarrow n + \pi^0$, $\Sigma^+ \rightarrow p + \pi^0$, $\Sigma^+ \rightarrow n + \pi^+$, $\Sigma^- \rightarrow n + \pi^-$), turned out to be *startlingly long*: 10^{-8} s for kaons and 10^{-10} s for hyperons. The fact that the decay schemes included no leptons suggested that these decays are associated with the strong interaction, in which case the lifetime of the particles must be about 10^{-22}-10^{-23} s.

An elegant solution to both problems was found by the American physicist M. Gell-Mann and the Japanese physicist K. Nishijima. They assumed that the long lifetime of kaons and hyperons is associated with the *conservation of some hitherto-unknown physical quantity* (just like the stability of the proton is associated with the conservation of baryon number and the stability of the electron, with the electric charge). So, another characteristic of the elementary particles appeared, and not without humour it was called the *strangeness*. A new conservation law was established that is valid for strong and the electromagnetic interactions: *the total strangeness of the mesons and the baryons involved in the process is conserved.*

Table 3 tabulates the values of strangeness S for various mesons (antimesons) and baryons (antibaryons). The strangeness of an antiparticle equals the strangeness of a respective particle with the opposite sign.

It follows from *strangeness conservation law* that in collision of a particle with zero strangeness, a lambda (or sigma) hyperon may only

be produced together with a kaon (the total strangeness of the kaon and the hyperon is zero). But the production of a xi hyperon must be accompanied by the production of two kaons (for example, $p + p \to p + \Xi^0 + K^0 + K^+$). Omega hyperons have been observed to be produced in a beam of negatively charged kaons: $K^- + p \to \Omega^- + K^0 + K^+$.

The long lifetime of kaons is accounted for by the fact that the kaon is the lightest particle with a nonzero strangeness. It cannot decay either due to the *strong* interaction, or due to the *electromagnetic* interaction

Table 3 Strangeness

	$S = -3$	$S = -2$	$S = -1$	$S = 0$	$S = +1$	$S = +2$	$S = +3$
Particles	Ω^-	Ξ^0 Ξ^-	Λ^0 $\Sigma^+\Sigma^0\Sigma^-$	π^+ p n	K^+ K^0		
				$\pi^0\eta^0$			
Antiparticles			$\bar{K}^0$ K^-	π^- $\bar{p}$ $\bar{n}$	$\bar{\Lambda}^0$ $\bar{\Sigma}^+\bar{\Sigma}^0\bar{\Sigma}^-$	$\bar{\Xi}^0\bar{\Xi}^-$	$\bar{\Omega}^-$

since there is *no particle to which it could transfer its strangeness*. The kaon has one possibility: to decay by *weak* interaction, since in such interactions strangeness *is not conserved*. So, the decays of the type $K^0 \to \pi^0 + \pi^0 + \pi^0$ or $K^+ \to \pi^+ + \pi^0$ are controlled, despite the absence of leptons, exactly by the *weak interaction*, which in turn predetermines the long lifetime of kaons.

The long lifetime of the lambda hyperon stems from the fact that this hyperon is the lightest baryon with nonzero strangeness. The decay of the lambda hyperons into kaons (or rather antikaons) is absolutely prohibited by the law of conservation of baryon number, and the decay into nucleons is prohibited by strangeness conservation. The observed decays $\Lambda^0 \to p + \pi^-$ and $\Lambda^0 \to n + \pi^0$ occur owing to the *weak interaction*, which does not conserve strangeness.

The charged sigma hyperons Σ^- and Σ^+, too, can only decay through the *weak interaction*. The sigma hyperon cannot decay into a lambda

hyperon and a pion, since the mass difference of a sigma and a lambda hyperon is smaller than the pion mass. In the case of the neutral sigma hyperon, a decay is possible that *conserves strangeness* (through the *electromagnetic* interaction): $\Sigma^0 \rightarrow \Lambda^0 + \gamma$. Therefore, the lifetime of a Σ^0-hyperon is shorter than 10^{-14} s.

An omega hyperon and xi hyperons decay into hyperons with a smaller mass:

$$\Xi^0 \rightarrow \Lambda^0 + \pi^0, \quad \Xi^- \rightarrow \Lambda^0 + \pi^-, \quad \Omega^- \rightarrow \Xi^0 + \pi^-,$$
$$\Omega^- \rightarrow \Lambda^- + \pi^0, \quad \Omega^- \rightarrow \Lambda^0 + K^-.$$

Since for the omega hyperon $S = -3$, and for xi hyperons $S = -2$, then in these processes, too, *strangeness is not conserved*, which predetermines their slow (*weak*) nature.

So far we do not know which principles of symmetry underlie the law of conservation of strangeness. There is no doubt, however, that strangeness conservation is one of the most important properties of the strong and the electromagnetic interactions, which accounts for the observed processes of interactions in the world of mesons and baryons. Specifically, it is of fundamental importance that strangeness does not conserve in weak interactions. If it were conserved not only in strong and electromagnetic interactions, but also in weak interactions (as, for instance, the electric charge, the electron, muon, and baryon numbers), then in addition to the electron and the proton *there would exist eight more* (!) *stable* subatomic particles with a nonzero rest mass: K^+, K^0, Λ^0, Σ^+, Σ^-, Ξ^0, Ξ^-, Ω^-. What structure would the atom have then would be anyone's guess.

Interactions and Conservations

As it has already been noted, the *highest symmetry* is inherent in processes occurring due to the *strong* interaction. For them we have ten conservation laws (Table 4): *energy, momentum, angular momentum, electric charge, baryon number, space, charge,* and *time parity, strangeness, isospin.* In principle, we can add two more conservation laws for the *electron* and *muon* numbers. True, in strong interactions these laws are fulfilled simply because there are no leptons, since the electron and muon numbers of all the components are zero.

Turning to *electromagnetic* interactions, symmetry *becomes lower – isospin* conservation is no longer valid. *Yet more marked reduction is observed when we go over to the weak interaction.* In the world of weak interactions we will have to forsake four conservations at once: *space* and *charge parity, strangeness,* and *isospin.* In some cases *time parity* is violated as well. This is somewhat compensated for by the conservation of *CPT-parity,* and in the majority of cases *combined parity* as well.

A Curious Formula

Gell-Mann and Nishijima turned their attention to a rather curious fact. It turns out that the electric charge Q of a particle (in terms of the ratio of

the particle's charge to the electron's charge), the isospin projection I_ζ, the baryon number B and the strangeness S are related by the following simple relationship (the *Gell-Mann-Nishijima formula*):

$$Q = I_\zeta + \frac{B+S}{2}.$$

The reader can easily prove this relationship for any meson or baryon.

Table 4

Interactions and Conservation Laws

Interactions	Conserved quantity					
	Energy. Momentum. Angular momentum. Electric charge	Baryon, muon, electron numbers	CPT-parity	Space parity, Charge parity	Strangeness	Isospin
Weak	+	+	+	−	−	−
Electromagnetic	+	+	+	+	+	−
Strong	+	+	+	+	+	+

For example, for Ξ^--hyperon: $Q = -1$, $I_\zeta = -1/2$, $B = 1$, $S = -2$. In this case, we have $-1 = -1/2 + \frac{1-2}{2}$.

The Gell-Mann-Nishijima formula related the four (seemingly different) physical characteristics for any meson or any baryon. The existence of such a relation suggests that there is a definite *internal completeness* of the established description of the properties of strongly interacting particles.

The Unitary Symmetry of Strong Interactions

Consider a system of coordinates in which the abscissa axis is the projection of isospin I_ζ, and the ordinate system is $Y = B + S$, a quantity called the *hypercharge*. On this plane we will position all the baryons with $s = 1/2$: p, n, Λ^0, Σ^-, Σ^0, Σ^+, Ξ^-, Ξ^0.

The *eight baryons* with spin 1/2 form a hexagon in the plane I_ζ, Y. At each vertex of the hexagon there lies one baryon, in the centre two

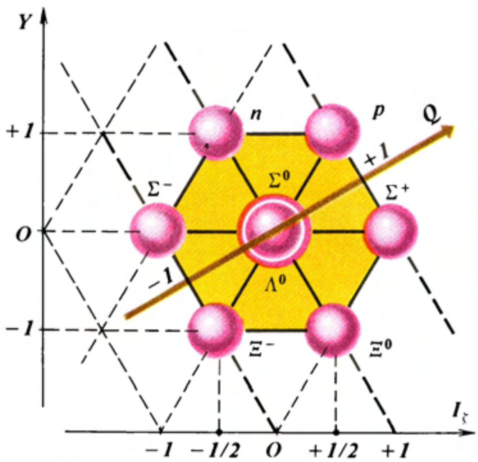


Fig. 118

baryons (Fig. 118). The arrangement of baryons in the plane allows the Q axis to be introduced.

Looking at Fig. 118, where all the eight baryons with spin 1/2 appear to be combined within a geometrically symmetric closed figure, one cannot ignore the fact that this is an example of some *concealed symmetry in nature*. This assumption is substantiated if we place other strongly interacting particles on the plane I_ζ, Y by combining them into groups with the same spin s. It appears that the eight particles with $s = 0$, which include all the mesons and antimesons (K^0, K^+, $\overline{K^0}$, K^-, π^+, π^0, π^-, η^0), form (on the plane I_ζ, Y) exactly the same hexagon as the eight baryons (Fig. 119). A result of no less interest is obtained for short-lived particles called *resonances*. Among these particles, which refer to baryons, we know nine particles with $s = 3/2$: Δ^-, Δ^0, Δ^+, Δ^{++}, Y_1^{*-}, Y_1^{*0}, Y_1^{*+}, Ξ^{*0}, Ξ^{*-}. In the plane I_ζ, Y they form a rectangle shown in Fig. 120, in which, however, one place is vacant – the vertex A. It is clearly seen in the figure that the missing particle (the missing baryon with spin 3/2) must be included in the isotopic singlet and have a negative charge and strangeness $S = -3$. You can imagine the satisfaction physicists derived when in 1964 the missing particle was actually found. So the hyperon Ω^-

was added to the list of elementary particles.

The *eight baryons*, the *eight mesons*, the *ten baryons* shown in Figs. 118–120 are called *supermultiplets*. Each supermultiplet contains several isotopic multiplets with different values of strangeness.

The symmetry that manifests itself through a union of mesons and baryons into several supermultiplets is the so-called *unitary symmetry*. The explanation of the mathematical nature of unitary symmetry lies beyond the scope of this book, and we only note that this symmetry

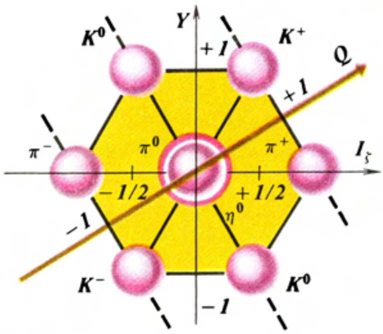


Fig. 119

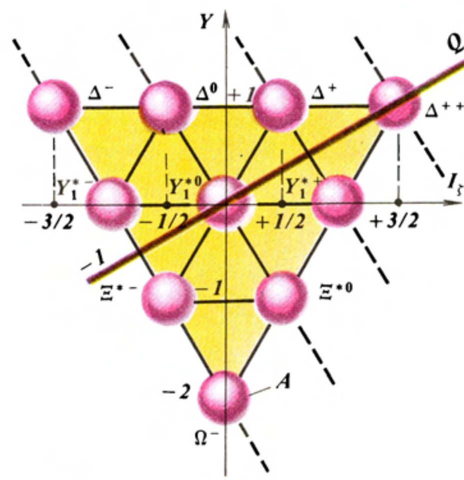


Fig. 120

establishes the internal relationship between the particles belonging to various isotopic multiplets and having different strangeness. The fact that a fairly numerous set of mesons and baryons (including resonances) can be compressed into a small number of eight-fold and ten-fold supermultiplets suggests that in the world of strongly interacting particles there exists a general order.

17 Quark-Lepton Symmetry

The symmetry between quarks and leptons today appears quite significant. It suggests that with all their striking dissimilarity, there is something common in the nature of these particles. It is the creation of the unified theory of quarks and leptons that, it seems, is going to be the main effort of physicists in future.

L. Okun

Up until recently physicists were bewildered by the disagreement between the abundance of hadrons and a modest number of lepton types. Perhaps that is why a hypothesis put forward in 1964 appeared to be so attractive. According to this hypothesis, all hadrons consist of several elementary "building blocks" called *quarks*. With the passage of years the quark hypothesis gradually gained ground. It is formulated as the rule: *the number of quark types must be equal to the number of lepton types*. This rule reflects the *quark-lepton symmetry*, which is as yet quite enigmatic.

Quarks

Unitary symmetry allows the existence of supermultiplets not only of eight or ten particles, but also other ones; specifically, supermultiplets are possible that contain only three particles. In the plane I_ζ, Y these "particles" form a triangle shown in Fig. 121*a*. The appropriate "antiparticles" form a triangle in Fig. 121*b*. In the figure, u, d, s stand for the "particles" and $\bar{u}$, $\bar{d}$, $\bar{s}$ for the "antiparticles".

Among the known elementary particles (antiparticles) there are no "particles" that are included in the triplets shown in Fig. 121. And yet these "particles", called *quarks*, for more than 15 years now have attracted the attention of physicists. In 1964 Gell-Mann and Zweig pointed out that three quarks in combination with three antiquarks can, in principle, be those "building blocks" of which all the known hadrons (mesons and baryons) and their antiparticles are constructed.

The characteristics of quarks u, d, s and their respective antiquarks are summarized in Table 5. Quarks do not have an integral, but fractional electric charge ($+2/3$ or $-1/3$). They are fermions (spin 1/2); this is only natural because only out of fermions can we construct both fermions and bosons (an odd number of fermions gives a fermion, an even number of fermions gives a boson). Quarks u and d have no strangeness, quark s has the strangeness $S = -1$ (the s-quark is, as it were, a carrier of strangeness).

Hadrons are constructed out of quarks according to the following simple rule: the *baryon consists of three quarks* (antibaryon out of three antiquarks), and the *meson, out of a quark and an antiquark*. So, for example, pion π^+ has the quark structure $u\bar{d}$, and its antiparticle (pion π^-), the structure $\bar{u}d$. The structure of kaons has the strange antiquark $\bar{s}$($K^+ = u\bar{s}$, $K^0 = d\bar{s}$). The quark structure of long-lived baryons is represented in Table 6. It is seen that the structure of most baryons

includes pairs of identical quarks, and in the hyperon Ω^- all the three quarks are identical. Furthermore, different baryons may have the same quark structure (hyperons Λ^0 and Σ^0). This means that a quark may be in different states. So we should take into account the possibility of two spin states of a quark. This does not need to be considered in the case of the hyperon Ω^-. The spin of this hyperon being 3/2, all three *s*-quarks are in the same spin state. Quarks are fermions, therefore, according to Pauli's exclusion principle, the three above *s*-quarks must differ in some

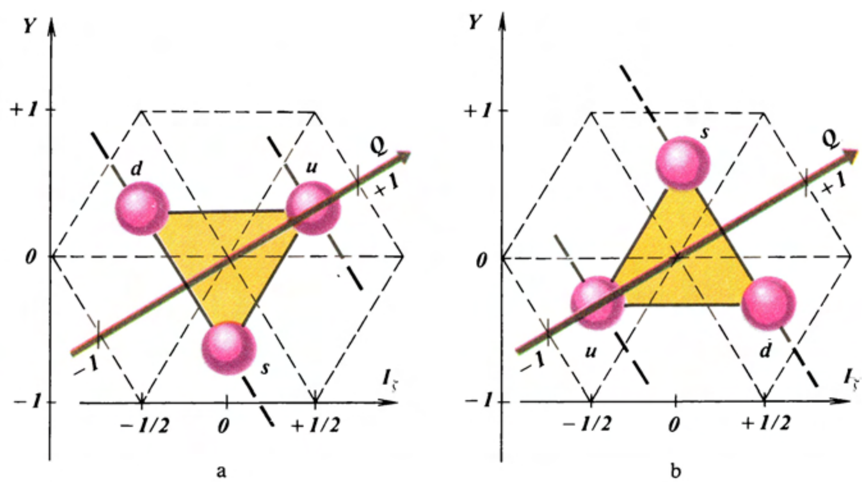


Fig. 121

additional parameter. In quark theory this parameter is called "colour".

According to modern views, each quark (antiquark) comes in *three varieties*, conditionally called *colours*. So, for example, there is a red *s*-quark, yellow *s*-quark, and blue *s*-quark. To be sure, the notion of quark colour should not be taken literally.

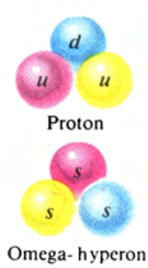


Fig. 122

Significantly, every baryon includes quarks of *different* colours (Fig. 122). Using the colour terminology, we can say that in each baryon three main colours are blended and so baryons can be viewed as *colourless* (white) objects. Mesons are colourless as well, since the colour of an antiquark is always complementary in relation to the quark colour in this meson. The theory of coloured quarks (quantum chromodynamics) explains why we do not encounter in nature particles consisting, say, of two or four quarks and, specifically, individual (free) quarks. This is related to the fact that the *hadrons* (*antihadrons*) *observed in nature must be colourless by all means*. Clearly, we cannot produce a colourless combination from one, two, or four quarks.

The theory shows quite convincingly why the hadrons occurring in nature must be colourless. But how stringent is the requirement of colourlessness? The final answer, clearly, comes from experiment. The

Table 5 Quarks and Antiquarks

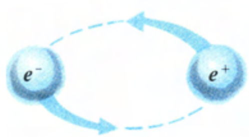
		Electric charge Q	Baryon number B	Isospin projection I_z	Strangeness S
Quarks	u	$+2/3$	$+1/3$	$+1/2$	0
	d	$-1/3$	$+1/3$	$-1/2$	0
	s	$-1/3$	$+1/3$	0	-1
Antiquarks	$\bar{u}$	$-2/3$	$-1/3$	$-1/2$	0
	$\bar{d}$	$+1/3$	$-1/3$	$+1/2$	0
	$\bar{s}$	$+1/3$	$-1/3$	0	$+1$

Table 6 Quark Structure of Baryons

p	n	Λ^0	Σ^+	Σ^0	Σ^-	Ξ^0	Ξ^-	Ω^-
uud	*udd*	*uds*	*uus*	*uds*	*dds*	*uss*	*dss*	*sss*

experimental search for free quarks has been under way for 15 years already, but to no avail.

Having read that no free quarks have as yet been found, the reader may doubt the physical reality of the quark hypothesis and may regard it as just an elegant mathematical trick. In 1965 Ya. Zeldovich wrote: "The dilemma facing physics now can be formulated as follows: either only the classification and symmetry properties of known particles are clarified, or this symmetry is a consequence of the existence of quarks, that is, absolutely new fundamental type of matter, atomism of new type". Another decade passed and physicists saw that the quark hypothesis was associated with the existence of a new type of atomism. In other words, by the end of the 1970s physicists no longer doubted that quarks in hadrons really exist.



Positronium

Charmonium

Fig. 123

What made them so sure? In the first place, three quarks (plus three antiquarks) made it possible to build all the hadrons (antihadrons) discovered up until 1974. It is remarkable that this construction produced no extraneous objects – all the particles constructed from quarks (antiquarks) according to the rules mentioned earlier have eventually been found experimentally. The quark model enabled the physicists to work out correctly the various characteristics of hadrons, the probabilities of transformations, and so on. Experiments on the scattering high-energy electrons at nucleons allowed quarks to be literally groped for within nucleons. Conclusive evidence for the validity of the quark hypothesis came from the discovery of new types of particles that were christened *charmed particles.*

The Charmed World

In November 1974 at the accelerator of Stanford, USA, a new particle of mass about 6000 m and lifetime about 10^{-20} s was discovered. This particle is now known as the *J/ψ-meson.* The spin of the J/ψ is unity. Like mesons π^0 and η^0, J/ψ is truly neutral.

The new meson does not conform to any of the earlier theoretical schemes, it would have a lifetime approximately 1000 times shorter. To describe the quark structure of J/ψ-meson a new quark had to be introduced – the so-called c-quark and a new conservable quantity called "charm". Just like strangeness and parity, charm is conserved in strong and electromagnetic interactions, but not in weak ones. It is *charm conservation* that is responsible for the relatively long lifetime of the J/ψ-meson.

When the c-quark was introduced, the number of quark types became equal to four. Note that the c-quark is the carrier of charm, just like the s-quark is the carrier of strangeness. The electric charge of the c-quark is $+2/3$.

The quark structure of the J/ψ-meson is $c\bar{c}$ (this structure explains specifically the true neutrality of the meson). The structure $c\bar{c}$ is called the *charmonium* and it is treated as an atom-like system resembling the long-known *positronium* (Fig. 123). Recall that the positronium represents an "atom" consisting of an electron and a positron orbiting

around a common centre of mass. (The existence of the bound system including an electron and its antiparticle was established experimentally in 1951; the lifetime of the positronium is as long as 10^{-7} s.) Like any atom, the charmonium is characterized by a system of energy levels. The J/ψ-meson corresponds to one of the levels of charmonium. Soon after the J/ψ-meson had been discovered, a number of mesons were found as well (ψ^1, χ_0, χ_1, χ_2, and so on), which could be correlated with various levels of charmonium. Studies of charmonium properties is of great interest – they provide information about the interaction of quarks.

The charms of the c-quark and $\bar{c}$-antiquark have opposite signs. Therefore, the resultant (total) charm of the $c\bar{c}$-structure is zero. It is said that the $c\bar{c}$-structure has a *hidden charm.* Mesons with *open charm* were discovered in the summer of 1976: the D^0-meson ($c\bar{u}$-structure) and the D^+-meson ($c\bar{d}$-structure). Their behaviour appeared to be in complete agreement with the hypothesis of the charmed c-quark. In 1977 the F^+-meson ($c\bar{s}$-structure) was discovered, which apart from charm also has strangeness.

The discovery of charmed particles proved experimentally the existence of the c-quark. And since the very c-quark and its properties are closely linked to those of u, d, s quarks, then the *quark model as a whole* was proved experimentally.

Returning to the issue of unitary symmetry of the strong interaction, note that allowing for charm the supermultiplets of hadrons take the form of *volume* bodies (polyhedrons) in the space in which the axes are I_ζ, Y, C (recall that here C is charm, I_ζ is the isospin projection, Y is the sum of the baryon number and strangeness called the *hypercharge*). The supermultiplets presented in Figs. 118-120 (see Chapter 16) are the sections of such polyhedrons by the plane $C = 0$. Given in Fig. 124 is an example of the polyhedron corresponding to the meson supermultiplet consisting of fifteen mesons.

Quark-Lepton Symmetry

It is to be stressed that the *representatives of hadrons in weak interactions are quarks.* Consider two examples: the decay of the neutron ($n \to p + e^- + \bar{\nu}_e$) and the collision of a neutrino with a neutron ($\nu_e + n \to p + e^-$). The decay of the neutron comes down to the decay of one of the d-quarks, which is a constituent part of the neutron, into a u-quark and leptons:

$$d \to u + e^- + \bar{\nu}_e.$$

The collision of a neutrino with a neutron resulting in the production of a proton and an electron boils down to the collision of a neutrino and a d-quark that enters the composition of the neutron, with the result that the d-quark turns into a u-quark and in the process an electron is produced:

$$\nu_e + d \to u + e^-.$$

The above processes involve a pair of $e\nu_e$ (or $e\bar{\nu}_e$) leptons and a pair of ud-quarks. Other forms of weak processes are also possible. So, for

example, the lepton pair μν may interact with the quark pair *us*. Any weak process is an interaction of a *lepton pair with a quark pair*.

This removes the discrepancy between the small number of lepton types and the vast number of hadrons. The number of leptons must be compared not with the number of hadrons but with the number of quarks. It appears that between leptons and quarks some sort of symmetry exists: *the number of lepton types must be exactly equal to the number of quark types.* This conclusion follows from the theory based on

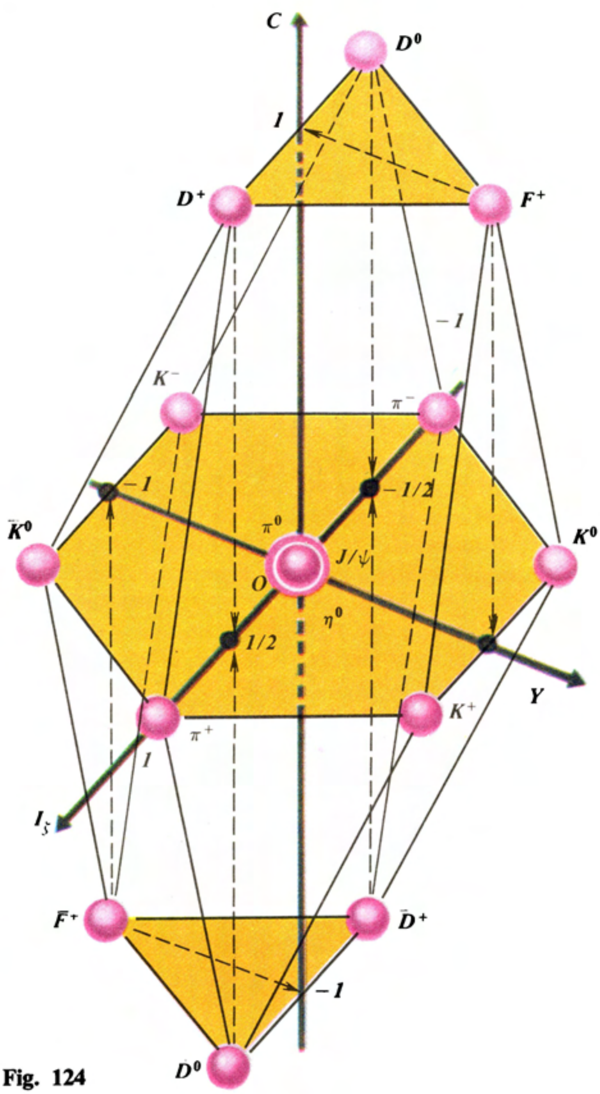


Fig. 124

a vast body of experimental evidence and, in particular, on the data on the decay of strange particles and the nonconservation of space parity in weak interactions.

Before the charmed particles had been discovered, there was no symmetry between leptons and quarks: only three quarks (*u*, *d*. *s*) corresponded to the four leptons e^-, ν_e, μ^-, ν_μ. Therefore, the *c*-quark was a welcome arrival.

The scheme of four leptons and four quarks suffered, however, from one drawback. The number of leptons (quarks) involved was insufficient to account for the nonconservation of combined parity in decays of neutral kaons (the nonconservation of combined parity was noted specifically in Chapter 14). At least six leptons (and as many quarks) were required.

A New Discovery

The first evidence for the existence of a fifth lepton came in 1975. In 1977 the hypothesis became certainty. The fifth lepton was called the *tauon* (τ^-). Its mass was found to be 3500 *m*. Apart from the new lepton another neutrino – tauon neutrino (ν_τ) – must exist.

In the summer of 1977 at the Fermi National Accelerator Laboratory, USA, superheavy mesons with a mass of about 20 000 *m* (epsilon-messon ϑ) were discovered. It was found that these mesons represented a structure of a quark and an antiquark of a new type. This quark (*b*-quark) is the carrier of the quantity that is conserved in strong interactions. This quantity was dubbed *beauty* – this explains the abbreviation for the fifth quark. The electric charge of the *b*-quark is $-1/3$, its mass is about 10 000 *m*.

At present, a search is under way for beautiful hadrons, and also for a sixth quark. If the third quark (*s*-quark) is called *strange*, the fourth quark (*c*-quark) is *charmed*, the fifth quark (*b*-quark) is *beautiful*, but it has been decided to call the sixth quark *true* quark, or *t*-quark.

The scheme of the six leptons and six quarks seems to be quite attractive to physicists today. The future will show whether or not this quark-lepton scheme is a final one or the number of leptons (quarks) will continue to grow.

A Conversation Between the Author and the Reader About the Role of Symmetry

In the 20th century the principle of symmetry encompasses an ever-increasing number of domains. From crystallography, solid-state physics, it expanded into the science of chemistry, molecular processes and atomic physics. It is beyond doubt that we will find its manifestations in the world of the electron, and quantum phenomena will obey it as well.

V. Vernadsky

...The thing that makes the world go round!

A. Pushkin

The Ubiquitous Symmetry

AUTHOR. In 1927 the prominent Soviet scientist V. Vernadsky wrote: "A new element in science is not the revelation of the principle of symmetry, but the revelation of its universal nature". I think that this book provides plenty of evidence to support the notion of the *universality of symmetry.*

READER. The universality of symmetry is startling. Symmetry establishes the internal relations between objects or phenomena which are not outwordly related in any way. The game of billiards and the stability of the electron, the decay of the neutron and the reflection in a mirror, a pattern and the structure of diamond, a snowflake and a flower, a helix and the DNA molecule, the superconductor and the laser....

AUTHOR. The ubiquitous nature of symmetry is not only in the fact that it can be found in a wide variety of objects and phenomena. The very *principle* of symmetry is also very general, without which in effect we cannot handle a single fundamental problem, be it the problem of life or the problem of contacts with extraterrestrial civilizations. Symmetry underlies the *theory of relativity, quantum mechanics, solid-state physics, atomic* and *nuclear physics, particle physics.* These principles have their most remarkable manifestations in the properties of the *invariance of the laws of nature.*

READER. It is quite obvious that here are involved not only the laws of physics, but also others, for example, the laws of biology.

AUTHOR. Exactly. An example of a biological conservation law is the *law of heredity.* It relies on the invariance of the biological properties when passing from one generation to the next. It is quite likely that without conservation laws (physical, biological and others) our world simply *could not exist.*

READER. Without energy conservation the world would be just a holocaust of random explosions associated with random appearances of energy from nothing.

AUTHOR. Imagine that on a nice day the laws of conservation of the electric charge and baryon number just cease to function. What would result then?

READER. Electrons and protons would then become unstable particles. And there would not be a single stable particle with a nonzero rest mass.

AUTHOR. You can easily imagine what sort of world we would live in. It would be a giant cluster of photons and neutrinos. Here and there some

11-1197

ephemeric formations would emerge only to decay quickly (in 10^{-10}-10^{-8} s) into some kind of photon-neutrino chaos.
And now imagine that suddenly the character of the symmetry of the wave electron function changes so that electrons become bosons.
READER. Maybe our world would then become a world of superconductivity? Electric current would pass along wires in this kind of world without resistance.
AUTHOR. Serious doubts emerge here in relation to wires themselves. Having stopped to obey Pauli's exclusion principle, electrons in all atoms would have to undergo a transition to the electron shell that is closest to the nucleus. The entire Periodic System would be thrown into disarray.
READER. Yes, indeed, symmetry permeates our world much deeper than it is apparent to our eyes.
AUTHOR. It took several centuries to conceive this, and the 20th century has been especially remarkable in this respect. As a result, the very concept of symmetry has undergone a substantial overhaul.

The Development of the Concept of Symmetry

AUTHOR. I hope that you understand how strongly today's picture of the *physically symmetrical world* differs from the *geometrically symmetrical cosmos* of the ancients. From antiquity to the present, the notion of symmetry has undergone a lengthy development. From a purely *geometrical* concept it turned into a fundamental notion lying at the foundation of *nature's laws*. We now know that symmetry is not only that which is visible to our eyes. Symmetry is not just *around us* but, moreover, it is *at the root of everything*.
READER. So it is not without good reason that the book is divided into two parts: "Symmetry Around Us" and "Symmetry at the Heart of Everything", isn't it?
AUTHOR. Yes, it is with good reason. The first part was devoted to geometrical symmetry. The second part was meant to show that the notion of symmetry is much deeper and that to grasp it one requires not so much a visual conception as thinking. As we pass on from the first to the second part we cover the path *from the symmetry of geometrical arrangements to the symmetry of physical phenomena*.
READER. As far as I understand, from the most general point of view the notion of symmetry is related to the *invariance under some transformations*. Invariance may be purely geometrical (the conservation of geometrical shape), but may have nothing to do with geometry, for example, the conservation of energy or biological properties. In exactly the same way transformations may be geometrical in nature (rotations, translations, commutations), and may be of another nature (the replacement of particles by antiparticles, the transition from one generation to the next).
AUTHOR. According to modern views, the concept of symmetry is characterized by a certain structure in which three factors are combined: (1) an *object* (*phenomenon*) whose symmetry is considered, (2)

transformations under which the symmetry is considered, (3) *invariance* (*unchangeability, conservation*) of some properties of an object that expresses the symmetry under consideration. Invariance exists not in general but only in as far as something is conserved.

READER. I heard that there exists a specially worked out *theory* of symmetry with its own mathematical tools.

AUTHOR. Such a theory does exist. It is called the *theory of groups of transformations*, or for short, just *groups theory*. The term "groups" was introduced by the father of this theory, the outstanding French mathematician Evariste Galois (1811-1832). These days any serious studies in the field of quantum physics, solid-state physics, and particle physics use the procedures of groups theory.

But back to the notion of symmetry. We will take a closer look at the fundamental nature of symmetry.

READER. This fundamentality is well grasped if one remembers that *symmetry limits the number of possible forms of natural structures*, and also *the number of possible forms of behaviour of various systems*. This has been repeatedly stressed in the book.

AUTHOR. It can be said that there are three stages in our cognition of the world. At the lowest stage are *phenomena*; at the second, the *laws of nature*, and lastly at the third stage, *symmetry principles. The laws of nature govern phenomena, and the principles of symmetry govern the laws of nature.* If the natural laws enable phenomena to be predicted, then the symmetry principles enable the laws of nature to be predicted. In the final analysis the predominance of symmetry principles can be judged from the actual presence of symmetry in everything that surrounds us.

READER. But perhaps there is some contradiction here with the fact that the most interesting discoveries in particle physics are related to *violations* of conservation laws, that is *violations of symmetry.*

AUTHOR. The violation of *P*-invariance in weak interactions is compensated for by the conservation of *CP*-parity. We can speak about the existence of a sort of *law of compensation for symmetry*: once symmetry is lowered at *one* level it is conserved at another, *larger* level. But this issue is worth considering in more detail. It is directly related to the problem of symmetry-asymmetry.

Symmetry-Asymmetry

AUTHOR. Faith in the primordial symmetry (harmony) of nature has always been an inspiration for scholars. And today it inspires them from time to time to undertake a search for a unified theory and universal equations.

READER. This search is not altogether without success. Suffice it to mention Einstein's theory of relativity or the discovery by Gell-Mann of unitary symmetry in strong interactions.

AUTHOR. That is all very well. But it should be noted that discoveries of new symmetries in the surrounding world *by no means brings us any nearer* to the cherished unified theory. The picture of the world, as we progress along the path of cognizing it, *becomes ever more complicated*,

and the very possibility of the existence of such universal equations becomes ever more doubtful. In the book *In Search of Harmony* by O. Moroz there is a rather graphic comment: "Physicists are chasing symmetry just as wanderers in the desert chase the elusive mirage. Now on the horizon a beautiful vision appears, but when you try and approach it, it disappears, leaving a feeling of frustration...". What is the matter here?

READER. Perhaps, it is that symmetry must be treated as no more than *ideal norm*, from which there are always *deviations* in reality.

AUTHOR. Right. But the problem of symmetry-asymmetry must be understood more deeply. Symmetry and asymmetry are so closely interlinked that they must be viewed as *two aspects of the same concept*. Our world is not just a symmetrical world. It is a symmetrical-asymmetrical world. The French poet Paul Valéry (1871-1945) was right in saying: "The world has orderly shapes strewn over it". He went on to observe that "events that are most astounding and most asymmetrical in the short run acquire a measure of regularity in the longer run".

READER. In the preliminary remarks it was stressed that the world exists owing to the *unity of symmetry and asymmetry*.

AUTHOR. The gist of the matter consists in that the unity of symmetry and asymmetry is the *unity of dialectically opposite categories*. It is similar, for example, to the unity of essence and phenomenon, necessity and chance, the possible and actual. The Soviet philosopher V. Gott in his book *Symmetry and Asymmetry* notes that "symmetry discloses its content and meaning through asymmetry, which in itself is a result of changes, or violations, of symmetry. Symmetry and asymmetry is one of the manifestations of the general law of dialectics – the law of unity and conflict of opposites".

Like two dialectically opposite categories, symmetry and asymmetry *cannot exist independently*. We have already said that in an *absolutely* symmetrical world you would observe nothing – no objects, no phenomena. In exactly the same way, an *absolutely* asymmetrical world is impossible too.

READER. It seems that the more we grasp the symmetry of nature, the more asymmetry comes out.

AUTHOR. Exactly. Therefore, any search for a unified theory or universal equations is bound to fail, as is any attempt to consider symmetry *separately from asymmetry*.

On the Role of Symmetry in the Scientific Quest for Knowledge

AUTHOR. Symmetry principles are extremely important in the great mystery called the *scientific quest for knowledge*. Any *scientific classification* is based on revealing the symmetry of the objects being classified. Objects or phenomena are grouped together by their common features that are conserved under some transformations.

A good example is the Periodic System of elements suggested by the great Russian chemist D. Mendeleev (1834-1907). From period to period is preserved the community of properties of elements, that enter a column of the table, for example lithium, sodium, potassium, rubidium, caesium. The behaviour of elements varies in the same way within a period for various periods.

READER. It seems to me that any classification is based not only on *symmetry*, but also on the *asymmetry* of properties.

AUTHOR. Right. It would make no sense to note the common properties of lithium, sodium, potassium, if these properties were also possessed by other elements within a period. The symmetry of properties of appropriate elements from various periods is only of significance in combination with the asymmetry of properties of elements within the same period. *Classification also presupposes both the conservation (community) and change (differences) of properties of the objects being classified.*

READER. Now I am beginning to understand the thesis about the dialectic unity of symmetry and asymmetry. This is the *unity of conservation and change*, and the *unity of community and difference.*

AUTHOR. Speaking about symmetry principles, we should always keep this unity in mind.

Symmetry thus lies at the root of all classifications. Crystals, for example, are classified by the type of symmetry of the crystalline lattice, by the properties of atomic binding forces, by electric and other properties. The classification of atoms is based on the community and differences in the structure of their emission spectra.

When dealing with an unknown object or phenomenon, one should above all identify the factors that *are conserved under given transformations.* Hermann Weyl noted that when one has to have something to do with some object having a structure, one should try and determine the transformations that leave the structural relations unchanged. You may hope that following this path you will be able to get a deep insight into the inner structure of the object.

By applying symmetry to the development of scientific classifications in structural studies, one can in the final analysis make *scientific predictions.* I think that some examples of such predictions are already known to you.

READER. For example, Mendeleev predicted a number of then-unknown chemical elements and gave a correct description of their properties. It is also worth noting Gell-Mann's prediction of the existence of omega hyperon.

AUTHOR. No less instructive is the example of the prediction of the *displacement current.* The outstanding British physicist James Clerk Maxwell (1831-1879) was keen enough to uncover in the phenomenon of electromagnetic induction discovered by Faraday the production of the alternating electric current by the alternating magnetic field. Having assumed that there exists also a *similar* inverse effect (the alternating magnetic field is produced by the alternating electric field), Maxwell put

forward the famous hypothesis of the displacement current, which then enabled him to formulate the laws of electromagnetism. Moroz writes in his book *In Search of Harmony*: "When we try to solve the enigma of what prompted Maxwell to that decisive step, what led him to the idea of the displacement current, the circumstances lead us to quite a probable answer–symmetry. The symmetry between electricity and magnetism. The fact that Maxwell noticed it could be that insight without which no great discovery can be made".

READER. Perhaps in the example of Maxwell one should speak not so much about *symmetry* as about *analogy*?

AUTHOR. The *method of analogy* is based on the principle of symmetry. It presupposes looking for common properties in different objects (phenomena) and the extension of this community to include other properties. Speaking about the role of symmetry in the process of scientific quest, we should pay special attention to the application of the method of analogy. In the words of the French mathematician D. Poia "there do not, perhaps, exist discoveries either in elementary or higher mathematics, or even any other field, which could be made without analogies".

READER. It seems to me that Poia has somewhat exaggerated the role of analogies.

AUTHOR. No more than the role of symmetry. The universality of the method of analogy, which has really been widely used in all the scientific disciplines *without exception*, is, in essence, the universality of the principles of symmetry. *Physical models* of objects (phenomena) are constructed exactly by the use of analogies. The DNA molecule is modeled as a screw. The spin of a particle is modeled as the angular momentum of a body that spins like a top about its axis. The collision of a photon with an electron in the Compton effect is modeled as the collision of billiard balls.

Analogies between different processes allow us to describe them using *general equations*. A simple example: the swinging of the common pendulum, oscillations of atoms in a molecule, oscillations of the electromagnetic field in an oscillatory circuit with a capacitance and an inductance are symmetrical (analogous) in the sense that all these processes are described by the same mathematical equation (the differential equation of simple harmonic motions). One equation is suitable for describing the process of radioactive decay, the process of discharging a charged capacitor, the variation of air density with altitude when there is no wind, the decrease in the intensity of a light beam propagating through a medium (the differential equation of exponential decay). The *unity of the mathematical nature* of the processes at hand, which makes it possible to regard them as analogous, points the presence of *deep symmetry*.

READER. It is just amazing how wide in its scope our talk about the role and place of symmetry appeared.

AUTHOR. It would have been much wider if we had included the questions connected not only with the *scientific* activities of man but also

with other aspects of his life, for example, *engineering, architecture, arts.*
READER. Symmetry is apparent in engineering and architecture. As regards painting, music, poetry, here, it seems, the dominance of symmetry is doubtful.

Symmetry in Creative Arts

AUTHOR. Note, above all, that the creative endeavour of man in all its manifestations *gravitate toward symmetry.* In his book *The Architecture*

Fig. 125

of the 20th Century the French architect Le Corbusier wrote: "Man needs order, without it all his actions lose their concordance, logical interplay. The more perfect is the order, the more comfortable and confident is man. He makes mental constructs on the basis of the order that is dictated to him by the needs of his psychology – this is the creative process. Creation is an act of ordering."
READER. These are the words of an *architect.* It is well known that the principles of symmetry are the governing principles for any architect. In

some cases the architect can do with the primitive symmetry of the rectangular parallelepiped, in other cases he uses a more refined symmetry, as, for example, in the building of the Council of Mutual Economic Cooperation in Moscow (Fig. 125).

AUTHOR. It would be better to speak not about "primitive" and "refined" symmetry but about how an architect solves the question of the *correlation between symmetry and asymmetry*. A structure that is asymmetrical on the whole may represent a harmonic composition out of

Fig. 126

symmetrical elements.

An example is the Cathedral of St. Basil the Blessed on Red Square in Moscow (Fig. 126). One cannot help admiring this bizarre composition of ten different parts, however the cathedral as a whole features neither mirror nor rotational symmetry. The architectural volumes of the cathedral, as it were, are superimposed on one another, intersecting, rising, competing with one another and culminating in a central cone. Everything is so full of harmony that it evokes the feeling of festivity. In

his book *Moscow*, M. Ilyin writes: "At first glance at the cathedral one can imagine that the number of architectural forms used in it is extremely large. It soon becomes clear, however, that the builders made use of only two architectural motifs – the eight-sided cylinder and the semicircle. The former defines the faceted forms of major volumes, whereas the second is represented by a wide variety of forms, from wide and graceful arches of the ground-floor to the peaked ogee gables."

READER. It turns out that the symmetry of the cathedral manifests itself in the repetition (conservation) of the two main motifs throughout the structure.

AUTHOR. Not just *conservation*, but also *variations*, or rather *development*, of them. The two main motifs do not simply recur in the various parts of the cathedral, but they seem to be developing as the glance of the observer traces the entire structure. This is a highly successful solution of the problem of symmetry-asymmetry. It is clear that without this startling asymmetry the cathedral would have lost much of its individuality.

READER. Obviously, it is impossible to calculate preliminarily so successful a solution of the problem symmetry-asymmetry. This is a work of genuine *art*. It comes from the genius of the builder, his artistic taste, his understanding of the beautiful.

AUTHOR. Quite so. It can be said that the *art* of architecture starts exactly when the architect hits upon the elegant, harmonic and original relationship between symmetry and asymmetry.

READER. True, in the modern massive construction of standard dwellings there is no question of the correlation between symmetry and asymmetry.

AUTHOR. These days this problem takes on a *new perspective*. It is now solved not on the level of a separate house, but on the level of a neighbourhood or even a town. In earlier times an architectural ensemble distinguished by its individuality was a house (temple, palace, etc.). But now more frequently this role is played by a group of houses, for example, a district. And it is on this level that present-day builders are called upon to solve the problem of symmetry-asymmetry.

READER. *Architecture* provides plenty of examples of the dialectical unity of symmetry and asymmetry.

AUTHOR. In *music* and *poetry* we have the same thing. In 1908 a Russian physicist, G. Vulf, wrote: "The heart of music – rhythm – consists in regular periodic repetition of parts of a musical piece. But regular repetition of similar parts to make up the whole constitutes the essence of symmetry. We can apply the concept of symmetry to musical piece all the more so since this piece is put down using notes, i. e., it acquires a spatial geometrical image, whose parts we can overview". He wrote elsewhere: "Just like musical pieces, symmetry may be inherent in literary works, especially poems".

READER. As for poems, you surely mean the symmetry of rhymes, stressed syllables, that is, in the final analysis *rhythm* again.

AUTHOR. Right indeed. But both in music and in poetry symmetry

cannot be reduced just to rhythm and cadence. Any good work (musical or literary) includes some *invariants of meaning* which permeate, varying their form, throughout the work. In his symphony a composer generally returns to the main theme many times, each time varying it.

READER. We have already seen something of the sort in the example of the Cathedral of St. Basil the Blessed.

AUTHOR. The preservation of a theme and its *varying* (*developing*) is here the *unity of symmetry and asymmetry*. And the more successful is an architect, a composer, or a poet in solving the problem of the correlation between symmetry and asymmetry, the higher is the *artistic value* of the produced piece.

Of direct concern to symmetry is *composition*. The great German poet Johann Wolfgang Goethe argued that "any composition is based on latent symmetry". To master the *laws of composition* is to master the laws of symmetry. The three principal laws of composition dictate the translation-identical repetition of elements of structure, contrasted repetition, varied repetition.

READER. This appears as an *ornament in time*.

AUTHOR. Really, temporal pattern of sorts. We will always admire the "patterns" produced by the great Russian poet A. Pushkin. Here is a relatively simple, elegant Pushkinian pattern:

That year was extraordinary,
The autumn seemed so loth to go;
Upon the third of January,
At last, by night, arrived the snow.
Tatyana, still an early riser,
Found a white picture to surprise her:
The courtyard white, a white parterre,
The roofs, the fence, all moulded fair:
The frost-work o'er the panes was twining;
The trees in wintry silver gleamed;
And in the court gay magpies screamed;
While winter's carpet, softly shining,
Upon the distant hills lay light,
And all she looked on glistened white.

(Tranlsated by Babette Deutsch)

We will not here examine the intonation in this excerpt. Let us reread it again and again to absorb the splendour of these poetical ornaments.

READER. Let us now turn to *painting*. Where does symmetry come in here?

AUTHOR. A picture is not a colour photograph, not by any means. The arrangement of figures, the combinations of stances and gestures, the countenances, the palette, the combination of shades, these all are carefully thought out in advance by the painter, whose main concern is to affect the beholder in the way he desires. By using asymmetrical elements, the painter strives to create something that on the whole features *latent symmetry*. The Russian painter V. Surikov wrote: "And how much time it

takes for the picture to settle down so that it would be impossible to change anything. The real sizes of each object are to be found. It is necessary to find the lock to piece together the elements. This is mathematics."

READER. To analyze the symmetry of an image it would be better perhaps to take a picture with a simple composition.

AUTHOR. We can take the picture *Madonna Litta* (Fig. 127) by the Italian genius Leonardo da Vinci.

Notice that the figures of the madonna and the child are inscribed into

Fig. 127

an *isosceles triangle*, which is especially clearly perceived by the beholder's eye. This immediately brings the mother and the child to the *forefront*. The madonna's head is placed absolutely exactly, but at the same time naturally, between the two symmetrical windows in the background. Through the window we can see the gentle horizontal lines of low hills and clouds. All of this adds to the air of the peace, which is intensified by the harmony of the blue colour with the yellowish and reddish tones.

READER. The inner symmetry of the picture is apparent. But what can be said about asymmetry?

AUTHOR. Asymmetry manifests itself, for example, in the body of the child, which cuts through the triangle. Furthermore, there is one highly expressive detail. Owing to the *closeness*, completeness of the lines of the madonna's figure an impression is produced that the madonna is totally indifferent to the surrounding world, and to the beholder too. The madonna is all concentrated on the child, she holds him gently and looks at him tenderly. She is thinking only about him. And suddenly all the closeness of the picture is gone, once we meet the stare of the child. And it is here that the inner poise of the composition is *disturbed*: the serene and attentive stare of the child is directed at the beholder, and so the picture *achieves its contact with the world*. Just try and remove that marvellous asymmetry, turn the face of the child to the mother, connect their glances. This will kill the expressiveness of the picture.

READER. It so happens that each time when we, marvelling at a piece of art, speak about harmony, beauty, emotional intensity, we thereby touch on the same inexhaustible problem – the problem of the correlation of symmetry and asymmetry.

AUTHOR. As a rule, in a museum or a concert hall we *are not concerned with that problem*. After all it is impossible *at the same time* to perceive and analyze the perception.

READER. The example of Leonardo's picture has convinced me that the analysis of symmetry-asymmetry is still useful: the picture begins to be perceived more acutely.

AUTHOR. We see thus that symmetry is dominant not only in the *process of scientific quest* but also in the *process of its sensual, emotional perception of the world*. Nature-science-art. In all of these we find the age-old competition of symmetry and asymmetry.

Literature

H. Weyl, *Symmetry* (Princeton, Princeton University Press, 1952)

M. Gardner, *The Ambidextrous Universe* (New York, Charles Scribner's Sons, 1979)

H. Lindner, *Das Bild der modernen Physik* (Leipzig, Urania-Verlag, 1975)

R. Feynman, *The Character of Physical Law* (London, Pergamon Press, 1965)

K. Ford, *The World of Elementary Particles* (New York, Blaisdell Publishing Company, 1963)

A. Kitaigorodsky, *Order and Disorder in the World of Atoms* (Moscow, Mir Publishers, 1980)

P. Davies, *The Forces of Nature* (Cambridge, Cambridge University Press, 1979)

E. Harrison, *Cosmology. The Science of the Universe* (Cambridge, Cambridge University Press, 1981)

TO THE READER

Mir Publishers welcome your comments on the content, translation and design of this book.

We would also be pleased to receive any other suggestions you may wish to make.

Our address is:

Mir Publishers
2, Pervy Rizhsky Pereulok
I-110, GSP, Moscow, 129820
USSR

THE FASTEST SPEED

by S. Filonovich, Cand. Sc. (Phys.-Math.)

The constant denoting the velocity of light in a vacuum is encountered in every branch of physics, and this universality brings out the unity of the physical world. Although it is now over three hundred years since the constant was first defined it yielded to scientific assaults only slowly, revealing as it did unexpectedly new phenomena. Its universality and the surprises that it threw up make any attempt to relate how the velocity of light came to be measured a minor history of physics. This book therefore explains the requisite science against the background of the historical personalities involved. Intended for teachers and school pupils.

10